الإحصاء والقياس الاجتماعي

الدكتور
مهدي محمد القصاص





دار نيبور
للطباعة والنشر والتوزيع

الإحصـــــاء
والقياس الاجتماعي

الإحصاء
والقياس الاجتماعي

د. مهدى محمد القصاص

أستاذ علم الاجتماع المشارك

كلية الآداب – جامعة المنصورة

2014



دار نيبور للطباعة والنشر والتوزيع ـ العراق



الطبعـة الاولـى
2014

رقم الايداع في دار الكتب والوثائق العراقية
رقم 565 لسنة 2014



دار نيبـور للطباعـة والنشـر والتوزيـع ـ العراق

العراق – ديوانية – شارع الرياضة
بغداد – شارع المتنبي
هاتف 009647702466027
هاتف 009647808994764
Dar_nippur@yahoo.com
Darnippur1@gmail.com

بسم الله الرحمن الرحيم

﴿وَكُلَّ شَيْءٍ أَحْصَيْنَاهُ كِتَابًا﴾

(النبأ: 29)

صدق الله العظيم

الفهرس

المقدمة

الإحصاء علم يهتم بالمعلومات والبيانات - ويهدف إلى تجميعها وتبويبها وتنظميها وتحليلها واستخلاص النتائج منها بل وتعميم نتائجها - واستخدامها في اتخاذ القرارات، وأدى التقدم المذهل في تكنولوجيا المعلومات واستخدام الحاسبات الآلية إلى مساعدة الدارسين والباحثين ومتخذي القرارات فى الوصول إلى درجات عالية ومستويات متقدمة من التحليل ووصف الواقع ومتابعته ثم إلى التنبؤ بالمستقبل.

ولم تعد البحوث الاقتصادية و الاجتماعية و الإدارية و غيرها في وقتنا المعاصر، و في ظل التقدم التكنولوجي الهائل في كافة ميادين حياتنا اليومية، تكتفي بمجرد عرض المشاكل و دراسة الظواهر و تحديد الأسباب و استخلاص النتائج و اتخاذ القرارات بطريقة سطحية مجردة عن أسلوب الإقناع و التقدير والقياس.

ولقد أصبح الاتجاه العام في مثل هذه البحوث و الدراسات هو استخدام طرق القياس الكمية ووسائل الإقناع الإحصائية و ذلك لتحديد الخصائص وإبراز الاتجاهات العامة في الظواهر الاجتماعية والإدارية، و تحليل العلاقات المتشابكة و المتبادلة بين الظواهر علي أساس موضوع غير متميز.

وعلم الإحصاء يعطي للباحثين في مجال العلوم الاقتصادية و الاجتماعية والإدارية، العديد من الطرق والأساليب اللازمة لضرورة القيام بالدراسات والبحوث الاقتصادية والاجتماعية و الإدارية والجغرافية علي أساس من القياس لحركة العديد من المتغيرات المحددة للظواهر موضوع الدراسة.

وتستخدم كلمة الإحصاء لتشير إلي عملية جمع البيانات الكمية و الأساليب المستعملة

في معالجة تلك البيانات، و قد نعني بهذه الكلمة أيضا عملية استخلاص بعض الاستنتاجات من دراسة عينة صغيرة لصياغة تعميمات يمكن تطبيقها علي مجتمعات اكبر حجما.

فبحوث الرأي العام علي سبيل المثال تقوم علي مقابلة و دراسة عينة صغيرة من أفراد المجتمع و لكن نتائجها تستخدم في الاستدلال علي اتجاهات الرأي العام في المجتمع ككل. و بذلك يمكن القول بان الإحصاء يشير إلي طرق تنظيم و تلخيص البيانات والي الأساليب التي تستخدم في تحليل و تفسير النتائج واستخلاصاتها يمكن تعميمها علي مجتمع الدراسة.

فالإحصاء هو علم يبحث فى طريق جمع الحقائق الخاصة بالظواهر العلمية الاجتماعية التى تتمثل فى حالات أو مشاهدات متعددة، وفى كيفية تسجيل هذه الحقائق فى صورة قياسية رقمية، وتلخيصها بطريقة يسهل بها معرفة اتجاهات الظواهر وعلاقات بعضها ببعض، ويبحث أيضاً فى دراسة هذه العلاقات والاتجاهات واستخدامها فى تفهم حقيقة الظواهر ومعرفة القوانين التى تسير تبعاً لها.

ومن هنا يتضح أن الإحصاء لا غنى عنه لأى باحث فى شتى المجالات المختلفة إذ اعتمد فى بحثه على الأسلوب العلمى. أي أن الإحصاء هو عصا الباحث التى تقوده إلى الطريق الصحيح، وهى الأداة التى تساعده على تفسير الظواهر التى يدرسها وتوضيح النتائج التى يحصل عليها ودلالات البيانات والأرقام التى يحصل عليها.

د. مهدى محمد القصاص

بيبان – كوم حمادة – البحيرة

2014

وصف المقرر وهدفه

يهدف هذا المقرر إلى تعريف طلاب قسم الاجتماع بعلم الإحصاء وأهميتها ودورها فى تسهيل عمل الباحث الاجتماعي فى التعامل مع مجتمع البحث بدءا من أخذ العينات وكيفية جدولة البيانات وتفريغها وتبويبها ووصفها (مقاييس النزعة المركزية والتشتت وأشكال توزيع البيانات) ودرجة ونوع العلاقات بين المتغيرات ومستوى قياسها ودلالتها واختباراتها كاختبار (ت، ف، كا2) الخ، وذلك بهدف إكساب الطالب مجموعة من الخبرات فى مجال الإحصاء الاجتماعي كى تساعده فى عرض نتائج البحوث الاجتماعية الكيفية بصورة كمية محدده وواضحة ومختصره ودقيقة.

وفيما يلي وصف المحتوى وهدف كل فصل حيث يهدف إلى تعريف وإفهام واستخدام الطالب لـ:

1 – التعريف بمعنى كلمة الإحصاء وتطور علم الإحصاء وأهمية الإحصاء للباحث الاجتماعي.

2 – أنواع المتغيرات المختلفة وكيفية التفرقة بين كل نوع منها وتصنيفها بشكل صحيح.

3 – العينات والمقصود بها وأنواعها المختلفة وطرق سحب العينات والطرق المختلفة لحساب حجم العينة من المجتمع المفتوح والمغلق.

4 – القدرة على تبويب البيانات الإحصائية التى يحصل عليها فى بحثه فى جداول تكرارية وأيضاً عرض هذه البيانات بالرسم البياني بطرقه المختلفة.

5 – القدرة على وصف وتحليل البيانات من خلال مقاييس النزعة المركزية المختلفة مثل الوسط الحسابي والوسيط والمنوال وتعريف الطالب بطرق حساب كل من تلك المقاييس السابقة من البيانات المبوبة والغير مبوبة وتدريب الطالب على تحديد نوع التواء التوزيع.

6 – القدرة على وصف البيانات من خلال مقاييس التشتت المختلفة مثل المدى والتباين والانحراف المعياري والانحراف المتوسط وتعريف الطالب بطرق حساب كل من تلك المقاييس السابقة من البيانات المبوبة والغير مبوبة.

7 – القدرة على تحليل التباين بين متغيرين أو أكثر عن طريق حساب قيمة نسبة «ف» ومقارنتها بقيمة «ف» الجدولية لتحديد مدى دلالتها إحصائيا.

8 – تمكين الطالب من القدرة على استخدام اختبار «ت» لتحديد ودراسة العلاقة بين متغيرين فقط متجانسين وغير متجانسين عن طريق حساب قيمة «ت» ومقارنتها بقيمة «ت» الجدولية لتحديد مدى دلالتها إحصائيا.

9 – القدرة على استخدام اختبار «كا2» لتحديد ودراسة العلاقة بين متغيرين عن طريق حساب قيمة «كا2» ومقارنتها بقيمة «كا2» الجدولية لتحديد مدى دلالتها إحصائيا.

10 – تمكين الطالب من القدرة على تقدير قوة العلاقات بين المتغيرات من خلال استخدام معاملات الارتباط المختلفة والتنبؤ بقيمة متغير عن طريق معرفة قيمة متغير آخر من خلال حساب معادلة خط الانحدار بين المتغيرين.

11 – تمكين الطالب من القدرة على تقدير ثبات وصدق الاختبار من خلال حساب قيمة معامل الثبات ومعامل الصدق.

ولتحقيق الهدف من ذلك المحتوى يستلزم استخدام بعض الوسائل منها:

1 – جهاز عرض الشفافيات.
2 – جهاز كمبيوتر.
3 – داتا شو.
4 – سبورة بيضاء وأقلام بألوان مختلفة.

ويتم قياس ذلك من خلال التقويم وفق الأسباب الآتية:

1 – مناقشات.
2 – أوراق عمل.
3 – مجموعات عمل لحل التمارين.
4 – الاختبار التحريري .

الفصل الأول

علم الإحصاء تعريفه و أهميته

أولا: تعريف علم الإحصاء.

ثانيا: أهمية علم الإحصاء.

ثالثا: تطور علم الإحصاء.

رابعا: علاقة علم الإحصاء بالعلوم الاجتماعية.

أولا: تعريف علم الإحصاء

هو فرع من فروع الرياضيات يشمل النظريات و الطرق الموجهة نحو جمع البيانات ووصف البيانات و الاستقراء و صنع القرارات[1].

و عندما نتكلم عن علم الإحصاء لا نعنى بذلك البيانات الإحصائية وإنما نقصد حينئذ الطريقة الإحصائية. وهى الطريقة التى تمكننا من جميع الحقائق عن الظواهر المختلفة فى صورة قياسية رقمية وعرضها بيانيا ووضعها فى جداول تلخيصية بطريقة تسهل تحليلها بهدف معرفة اتجاهات هذه الظواهر وعلاقات بعضها ببعض[2].

ولقد كان الهدف الرئيسى من علم الإحصاء قديما هو عد أو حصر الأشياء المراد توفير بيانات إحصائية عنها، وكانت الجهة التى تقوم بإعداد الإحصاءات على مستوى الدولة تعرف بمصلحة التعداد ولذلك كان التعريف القديم لعلم الإحصاء أنه علم العد، أي العلم الذي يشتمل على أساليب جمع البيانات الكمية عن المتغيرات والظواهر موضوع الدراسة.

ولكن مع تطور المجتمعات وتشابه جوانب الحياة الاقتصادية والاجتماعية الحديثة بها، لم يعد مجرد توفير البيانات الكمية عن المتغيرات والظواهر موضوع الدراسة يفى بحاجات متخذى القرارات وصانعى السياسة العامة إلى تكوين صورة متكاملة الجوانب عن مجتمعهم والمجتمعات المحيطة به. فقام العلماء بتحديث نظريات علم الإحصاء وأساليبه وأدواته لكى يعين الباحثين وغيرهم على استخلاص استنتاجات معينة من البيانات الكمية التى أمكن لهم جمعها عن طريق العد.

(1) مصطفى زايد، الإحصاء ووصف البيانات، 1989، ص 23.

(2) فاروق عبد العظيم، مختار الهانسى، محمد على محمد، مبادئ الإحصاء، دار المعرفة الجامعية، ص 3.

من ذلك على سبيل المثال، أن نظرية العينات ساعدت الباحثين على استخلاص استنتاجات عديدة من دراسة عدد صغير من الأفراد أو الأشياء – العينة – وتعميم تلك الاستنتاجات على المجتمع الذى سحبت منه العينة بأسره ولذلك يعرف علم الإحصاء حديثاً بأنه: (علم متكامل يتضمن الأسلوب العلمى الضرورى لتقصى حقائق الظواهر واستخلاص النتائج عنها، كما يتضمن أيضاً أيضا النظرية اللازمة للقياس واتخاذ القرار فى كافة الميادين الاقتصادية والاجتماعية والسياسية والعسكرية)[1]

(1) حسن محمد حسن، مبادئ الإحصاء الإجتماعي، دار المعرفة الجامعية،2000، ص ص 15 – 16.

ثانيا: أهميه علم الإحصاء

لقد أصبح لعلم الإحصاء أهميه بالغه فى حياتنا الحديثة فصارت الإحصاءات مألوفة لدينا وتمثل جانبا مهما من المعلومات التى نطالعها كل يوم مثل جداول النقاط التى تحرزها أنديه كره القدم وتنشر فى الصحف والمجلات والتقديرات الخاصة بالتنبؤات الجوية ومؤشرات البورصة وانجازات الحكومة فى مجال الإسكان والتعمير والتغيرات التي تطرأ على أسعار العملات وأثمان السلع. وربما يتساءل المرء عن أهميه الإحصاء بالنسبة لدارس علم الاجتماع أو علم النفس معتقداً أن الإحصاء موضوع يدخل فى صميم تخصص التجاريين والاقتصاديين والواقع أن الباحث الاجتماعي والمتخصص فى العلوم الاجتماعية بوجه عام يحتاج فى كثير من الأحيان إلى استخدام الأرقام لكي يلخص ويعرض بها مجموعه من المشاهدات التى تتعلق بظاهرة يهتم بدراستها، فقد يطلب منه أن يقدم تقريراً عن مدي التطور الذى حققه برنامج معين لمحو الأمية بين نزلاء المؤسسة التي يعمل بها، وقد يكلف بدراسة الأسباب التى تجعل الذكور أكثر تقدما وحرصا على التعليم من الإناث فى المدرسة التى يشتغل فيها.

ففي كل مناسبة من هذه المناسبات سيحتاج الباحث أو الدارس إلى أداة من الأدوات الإحصائية لكي يستخدمها فى تلخيص أفكاره والتعبير عنها بصورة محدده ومؤثرة، فالعبارة التى مؤداها « لقد نجحنا فى محو أميه 90٪ من العاملين الأميين بالمصنع » أقوى وأشد من العبارة التى مفادها: « لقد نجحنا فى محو أمية عدد كبير من العاملين الأميين بالمصنع[1] يحتل الإحصاء (أو الأساليب الإحصائية) أهميه خاصة فى الأبحاث العلمية الحديثة، إذ لا تخلو أى دراسة أو بحث من دراسة تحليليه إحصائية تتعرض لأصل

(1) حسن محمد حسن، مرجع سابق، ص 16.

الظاهرة أو الظاهرات المدروسة فتصور واقعها فى قالب رقمى، وتنتهي إلى ابرز اتجاهاتها وعلاقاتها بالظاهرات الأخرى[1].

إن دراسة الإحصاء أمر له فوائد كثيرة بالنسبة لدارسي العلوم الاجتماعية وخاصة بعد أن تفتحت أمامهم مجالات عمل كثيرة فى تنظيمات الشرطة والعلاقات العامة بالشركات ومراكز البحوث وغير ذلك من مجالات العمل المختلفة. بل إن المعرفة بالإحصاء قد تفيد الإنسان على المستوى الشخصي فتكسبه مهارة التخطيط لحياته الاقتصادية الخاصة.

ولكن ينبغي أن نشير إلى أن النتائج التى تسفر عن تطبيق أداة إحصائية أو أكثر ليست نتائج قطعيه أو غير قابله للتمحيص والمراجعة. فإذا كانت الأدوات الإحصائية تستطيع أن تعين المرء على وصف البيانات وتصميم التجارب وعلى اختبار العلاقات بين الأشياء والوقائع التى يهتم بها إلا أن ذلك لا يلغى بصيرته السوسيولوجية وخبرته المهنية.

وبعبارة أخري، يقتصر دور الأدوات الإحصائية على توفير المؤشرات المبدئية التى تساعد الباحث على رفض أو قبول الفروض التى يقوم بدراستها في حدود درجه معينه من الثقة. والإحصاء أيضا أداه لا تستخدم إلا في العثور على إجابات عن أسئلة تتصل ببيانات يمكن التعبير عنها بصيغ كميه. وهناك فى مجال العلوم الاجتماعية موضوعات لا حصر لها لا يمكن صياغة البيانات الخاصة بها فى صورة كميه على نحو دقيق، ومن ثم لا يستطيع الباحث استخدام التحليل الإحصائي فى دراستها.

من ذلك على سبيل المثال، دراسة التجربة الدينية بين جماعه المؤمنين بدين معين، إذ أن عدد مرات تردد المرء على المسجد أو على الكنيسة في الشهر ليس دليلا فى حد ذاته على انه من الصالحين، ولكنه مؤشر مبدئي على الصلاح.

ومما يعكس أهميه علم الإحصاء أنها يستخدم فى توجيه عمليه جمع البيانات وفى تفسير العلاقات التى تعكسها تلك البيانات. ومن ابرز المجالات التى تستخدم فيها المعالجات

(1) فتحى عبد العزيز أبو راضى، مبادئ الإحصاء الإجتماعى، دار المعرفة الجامعية، ص 4.

الإحصائية إجراء المقارنة بين عديد من الأشياء فى كثير من المناسبات. ويمكننا القول أن الحياة الإنسانية سلسله من المواقف التى يتخذ فيها الفرد قراره بناء على ما تسفر عنه المقارنة التى يجريها بين عديد من الاحتمالات وهذه المقارنة فى جوهرها عمليه إحصائية تقترن بالقياس والتقييم والتقدير. فنجاح الإنسان فى حياته يتحدد وفق مقياس معين فى ذهنه يقدر به هذا النجاح، وحرية الفرد فى مجتمعه تقاس أيضا وفق معايير يتعارف عليها الأفراد فى مجتمعهم.

وبعبارة أخرى، إن حياتنا تذخر بعمليات من القياس والتقدير الإحصائي فنحن علي سبيل المثال،عندما ننزل إلى السوق لشراء سلعه معينه، فى موسم التنزيلات، نهتم وبطريقه لا شعورية بحساب ثمن هذه السلعة بالنسبة إلى إجمالي النقود التي فى حوزتنا ونقدر ما إذا كان الباقي من هذه النقود وسوف يكفينا حتى نهاية الشهر أم لا وما إذ كانت نسبة التنزيلات على السلعة حقيقية أو مزيفة 00000الخ فى كل هذه العمليات الفكرية نحن نستعين بعمليات إحصائية ومقارنات مستمرة بين المواقف المختلفة. فضلا عن ذلك، إن ما نطلق عليه ظاهرة اجتماعيه أو طبيعيه ما هو فى الواقع إلا سلسله متكررة من الواقع التى يمكن رصد حدوثها المستمر عبر فترة من الزمن وبنفس الوتيرة بطريقه إحصائية[1].

(1) حسن محمد حسن، مرجع سابق، ص ص 17 – 18.

ثالثا: تطور علم الإحصاء

تطور علم الإحصاء وتطبيقاته عبر سنوات طويلة، وتم ذلك بجهود كثيرة من العلماء من دول مختلفة وكان. التطور بطيئا إلى أن جاء القرن العشرين ليشهد معدلا هائلا للتطور فى النظريات الإحصائية فى مجالات كثيرة.

ويرجع الاهتمام بالإحصاء إلى عصور قديمه، وان تعداد السكان عند القدماء المصرين وفي الصين أمثلة توضح اهتمام الحكومات منذ القدم بالمعلومات الاجتماعية وذلك لأغراض التنظيم والتخطيط فى أحوال السلم والحرب.

ويبدو أن كلمه إحصاء (statistics) قد ظهرت لأول مره عام1749 وهى مشتقه من الكلمة اللاتينية (status) أو الايطالية (statista) وتعني كلاهما الدولة السياسية. ومن الطبيعي أن تكون الدولة أول من اهتم بجمع البيانات وذلك لإدارة شؤون البلاد خاصة عن السكان لأغراض حربية وضريبية، وامتدت بعد ذلك لتشمل إحصاءات حجم السكان والمواليد والوفيات والإنتاج والاستهلاك والثروة 00000الخ. وهكذا بداء العلم وتطوره باعتباره علم الدولة أو علم الملوك[1].

ولقد تطور علم الإحصاء من مجرد فكره الحصر والعد إلى أن أصبح الآن علما له قواعده ونظرياته ويرجع الفضل فى ذلك إلى كثير من العلماء من أمثال عائله برونلي Bernoulli وفردريك جاوس F.gauss وكيتليه Quetlet وجولتونF.galton وأخيرا كارل بيرسون Karl.pearson وبولي A.bowley وبول U.yule فيشر L.fisher و........الخ[2].

(1) مصطفى زايد، مرجع سابق، ص 19.

(2) فاروق عبد العظيم وآخرون، مرجع سابق، ص 3.

وجاء التطور في علم الإحصاء بصفه عامه ملازما وموازيا للتطور في نظرية الاحتمالات. فقد نشأت نظريه الاحتمالات على أساس رياضي فى (1494) بواسطة باسيولي Lucapacidi. ومن الدراسات الفلكية لكل من كبلر (1517 1630-) Keplr وجاليليو (1564 1642-) Galilio قاما بتطوير نماذج الاحتمالات. غير أن التاريخ الحقيقي لنظريه الاحتمالات بدء فى القرن السابع عشر حيث وضعت أسسها في عام 1654 بواسطة كلا من العالمين: باسكان (1623 1662 .Pascal.B) عالم الرياضيات والفيزياء والفيلسوف الفرنسى - وكذا العالم فرمات (1608 – 1665 Fermat).

وبعد ذلك بثلاث سنوات قام هينجينز (1629 – 1695 Huygens) بنشر كتيب صغير في موضوع المعالجة الرياضية لفرص الفوز في مباريات ورق اللعب وزهرة النرد.

وفى نفس الوقت تقريبا قام جرونت (1620 – 1674 grunt) بنشر ملاحظاته عن معالجة البيانات المتعلقة بالحكومة خاصة في النواحي الطبيعية والسياسية والتجارية والنمو والوفيات والأمراض.

وقد كان العمل الذى قام به هيجيتر دافعا للكثيرين لدراسة النظريات والمشاكل المتعلقة بمباريات الصدفة ومنهم برنوللى (1654 – 1705) ودي موافر – De Moivre (1667 1754) واربوثنوت Arbuthnott ولابلاس (1749 – 1827 laplace) وجاوس Gauss (1777 – 1855)[1]. ويعد العالم البلجيكى كتيليه (1796 – 1874) أول من وضع قواعد محددة لعلم الإحصاء، وكلمة إحصاء في الوقت الحاضر ذات معان متعددة فمنها يفهم جمع المعلومات التى تبين الحالة في الدولة مثل عدد المواليد والوفيات وبيانات عن المحاصيل والتجارة الخارجية ...... الخ ويسمى نشر الأجهزة الحكومية لمثل هذه المعلومات في شكل كتب وتقارير (بالإحصاء الرسمي).

وأخيرا يفهم بالإحصاء فرع من العلم له نظريته الخاصة. وعلم الإحصاء، شأنه في

---

(1) مصطفى زايد، مرجع سابق، ص ص 19 – 20.

ذلك شأن أى فرع آخر من فروع العلم له أسلوبه وموضوعات البحث الخاص به[1]

وكلمة إحصاء (Statistics) لها ثلاث معاني:

(1) **الإحصاءات أو البيانات:** مثال ذلك إحصاءات السكان والمواليد والوفيات والإنتاج – الصادرات – الاستهلاك.

(2) المؤشرات المحسوبة من عينة (العينة هى مجموعة جزئية من الوحدات محل الدراسة)

(3) علم الإحصاء: وهو فرع من فروع الرياضيات يشمل النظريات والطرق الموجهة نحو جميع البيانات ووصف البيانات والاستقراء وصنع القرارات[2].

ولقد تطور علم الإحصاء وتنوعت طرائقه، وأصبح له من القواعد ما يمكنه من القيام كعلم مستقل يمكن الاستعانة به في رسم وتحديد السياسات الاجتماعية التى ينتهجها المجتمع. كما برز دور الإحصاء – بما يقدمه من بيانات وإحصاءات – فى عمليات التخطيط والتنمية التي تمر بها مجتمعاتنا اليوم[3] ويمكن القول أن الإحصاء تخدم الباحثين في جميع الميادين العلمية وصانعى القرارات في شتي المجالات العملية، ولا يكاد يخلو ميدان من ميادين البحث العلمى إلا وطرقته الإحصاء وساهمت فيه مساهمة فعالة. وقد أثار روبرت بارسوز فى مستهل كتابه « التحليل الإحصائي » أن كلمة إحصاء لها أكثر من استخدام إلا أن أكثر الاستخدامات شيوعاً هو ذلك الذى يرى أن كلمة إحصاء تشير إلى تلك الأساليب والإجراءات التحليلية المستخدمة في معالجة البيانات الرقمية.

بمعنى أنه للحصول علي معلومات ذات قيمة من تلك البيانات الرقمية فإنها يجب أن

---

(1) غريب محمد سيد أحمد، الإحصاء والقياس فى البحث الإجتماعى، دار المعرفة الجامعية، 1989، ص 12.

(2) مصطفى زايد، مرجع سابق، ص 23.

(3) غريب محمد سيد أحمد، مرجع سابق، ص 13.

تخضع للتحليل الاحصائىStatistical Analysis بمساعدة تلك الأساليب والإجراءات والأدوات التي توفرها لنا الإحصاء.

ويذهب كل من Whittaker .Startup إلى وجود ثلاثة استخدامات لكلمة إحصاء.

أ – للإشارة إلى الحقائق الرقمية التى جمعت بطريقة منتظمة من الواقع الاجتماعي.

ب – تشير إلى الأساليب المستخدمة فى جمع، وتصنيف وتحليل البيانات الرقمية.

ج – للإشارة إلى صفة أو خاصية للعينة تحت الدراسة.

والقاموس الحديث لعلم الاجتماع الذى وضعه كل من George and Achilles Theocorson يقدم رؤية لا تختلف عما سبق فيما يتعلق بكلمة إحصاء سواء من حيث المعنى أو الاستخدام فهى تعني مجموعة من الأساليب التي تستخدم في جمع، وتصنيف، وتبويب وعرض وتحليل البيانات الكمية، والإحصاء بهذا المعني لا تقف عند حد الوصف Description بل تتعداه إلي مرحلة الاستنباطInduction والاستدلال Inference كما تستخدم كلمة إحصاء للإشارة إلي البيانات الرقمية والتى عادة ما تسمي إحصاءات حيث تأخذ صيغة الجمع.

ومن هنا فان كلمة إحصاء تعني تلك الأساليب والأدوات والإجراءات الإحصائية التى يلجأ إليها الباحث وهو بصدد القيام بدراسة ما فى عملية الجمع، وتصنيف، وتلخيص وعرض، و تحليل البيانات الرقمية[1].

---

(1) اعتماد علام، يسرى رسلان، أساسيات الاحصاء الإجتماعى، دار الثقافة للنشر والتوزيع، ص ص 7 – 8.

رابعا: علاقة علم الإحصاء بالعلوم الاجتماعية

تأثرت العلوم الاجتماعية وخاصة علم الاجتماع وعلم النفس وعلم السياسة بالتطورات. التي حققها علم الإحصاء، واستعان العلماء الاجتماعيون بمنهج جديد فى دراساتهم. وهو المنهج الاحصائى الذى ينطوي علي نفس خطوات المنهج العلمي في البحث، حيث يقدم علي عمليتين منطقيتين هما القياس و الاستنتاج، وإذن يقوم العالم بملاحظة الحقائق في البداية ثم يجري تجاربه ويرصد عددا من النتائج التي يستخلصها من تلك التجارب بنمط أو إطار عام للظاهرة. وبعد أن يقوم بصياغة نظريته علي ذلك النحو، ينتقل إلي عملية الاستنتاج التي تعينه علي التنبؤ بسلسلة من النتائج الأخرى.

ومن أشهر الدراسات السوسيولوجية التي اعتمدت علي المصادر الإحصائية، دراسة دور كايم عن الانتحار. وفيها يذهب إلى (انه إذا كان المرء يريد أن يعرف كل ما يتفرع عن الانتحار كظاهرة جمعيه فانه ينبغي أن ينظر إليها فى شكلها الجمعي من خلال البيانات الإحصائية) وقد اعتبر دور كايم أن المؤشرات الإحصائية عن الأسباب التى دفعت الأفراد إلى الانتحار بمثابة مصدر لمعرفه الدوافع المفترضة وراء الإقدام عليه. وهكذا نجد أنه قد وضع فروضه على أساس من الأرقام والإحصاءات التى رأي أنها تعين لنا اقرب نقطة لبدء بحثنا السوسيولوجى.

وقد حقق المنهج الإحصائي فى السنوات الأخيرة تقدما هائلا، وخاصة بعد استخدام الحاسبات الالكترونية، وذلك فى ميادين العلوم الاجتماعية المختلفة، وقد انعكس هذا التقدم بدورة على التطورات والأدوات الإحصائية ذاتها.

وقد استفاد علماء الاجتماع من المنهج الإحصائي في تطوير أدوات بحثهم وخاصة الاستبيان مما أمكنهم من دراسة آلاف المبحوثين فى فتره زمنية وجيزة، وتوافرت لدي الباحثين إمكانية اختبار العلاقة بين ما يرصدونه من ظواهر على أرض الواقع وما يفترضونه من افتراضات يحاولون بها تفسير ذلك الوقع.

وقد ساعد علم الإحصاء علماء السياسة على اقتحام مجالات عديدة من البحث السياسي مثل دراسة أنماط المشاركة السياسية وتكوين الرأي العام والحركات والتنظيمات السياسية. فلو أن عالم السياسة افترض أن هناك ثمة ارتباط بين مستوي تعليم الأفراد وتعليم من أدلوا بأصواتهم في الانتخابات فان البيانات التى يتسنى له الحصول عليها من الواقع عن مشاركه الأفراد فى التصويت الانتخابي وعن مستوياتهم التعليمية لا تنعقد المقارنة بينها إلا باستخدام المقاييس الإحصائية التى تكشف عن قوة الارتباط بين الميل للتصويت فى الانتخابات والمستوي التعليمي للأفراد. وبدون هذه المقاييس الإحصائية تظل البيانات والمعلومات الميدانية المتوافرة لدي الباحث بلا قيمه حقيقية.

ويستخدم علماء النفس الأدوات والأساليب الإحصائية أكثر من غيرهم فى القياس النفسى. ويعد علم النفس التجريبي وعلم النفس الاكلنيكي وعلم نفس الفروق الفردية من المجالات التى تعتمد اعتمادا جوهريا على المنهج الاحصائى فى تناولها لموضوعات الدراسة.

ومن يقرأ مرجعا فى القياس النفسي يجد أن علماء النفس يذهبون إلي أن كل شيء فى مجال علمهم قابل للقياس تقريباً فنجد لديهم مقاييس للذكاء وللشخصية وللعواطف والميول وللاضطرابات النفسية والأمراض العقلية وكل مقياس من هذه المقاييس يخضع، فى واقع الأمر لأساليب إحصائية صارمة تحدد مدى ثباته وصدقه فى قياس ما صمم لقياسه ويستخدم في المقارنة بين النتائج التى يتم التوصل إليها من دراسة عينه محدده من الأفراد وتلك التى يتم التوصل إليها من دراسة عينه أخرى[1].

(1) حسن محمد حسن، مرجع سابق، ص ص 18 20-.

وقد ظهر اهتمام كبير بتطبيق النظريات والطرق الإحصائية فى العلوم الاجتماعية، فقد أوضح كيتيليه (1796 – 1874) عالم الفلك الاجتماعي البلجيكي إمكان استخدام الاحتمالات والإحصاء لوصف وتفسير الظواهر الاجتماعية والاقتصادية وقدم مساهمات هامه فى الطرق الإحصائية فى تنظيم وإدارة الإحصاءات الرسمية – وقدم كذلك طريقه عامه للقياس فى الانثروبولوجيا – وقد ساهم عالم النفس الانجليزي جالتون (Galton 1911– 1822) في تطبيق الطرق الإحصائية في علم النفس، ووضع أساس علم القياس النفسى (psychometrics) وبدأ دراسة موضوع الارتباط والانحدار الذي اهتم به وطوره بعد ذلك عالم الإحصاء الانجليزي كارل بيرسون (Pearson 1857 –1936). بالإضافة إلي مساهمات أخري هامه.

كما قدم سبيرمان (Spearman 1863 –1945) عالم النفس الإنجليزي مساهمات فعاله في دراسة الارتباط ويعد من الرواد في دراسة وتطوير التحليل العاملي.

وقدم عالم الإحصاء الانجليزي جولست(1876 1937–) Gosset مساهمات هامه في مجال التحليل الإحصائي وخاصة في تفسير البيانات المتعلقة بالعينات كما يعد من الرواد المهتمين بتحليل نتائج العينات الصغيرة. وخلال الفترة السابقة كان الاهتمام كله مركزا علي المفهوم الكلاسيكي للاحتمال.

إن مفهوم التكرار النسبي لم يظهر بصوره ملموسة إلا في بداية القرن العشرين حيث تم صياغتها وظهورها في إطار منطقي بمعرفة فون مايسيس vonmises.

وعلي الرغم من أن الرواد من علماء الإحصاء كان إهتمامهم بوظيفة الاستقراء فان الجانب الأعظم من النظرية الإحصائية تم اكتشافه بعد عام 1920 تقريبا فمنذ مطلع القرن العشرين كان الاهتمام منصباً علي تطبيق الإحصاء علي مشاكل علوم الحياة وعلي التجارب الزراعية والصناعية.

كما أن العمل في هذه المرحلة كان مكثفا ومركزا علي التحليل الإحصائي وأساسه المنطقي،

وتمخض عن ذلك مساهمات قدمها عالم الإحصاء الانجليزي فيشر 1962– 1890 Fisher ومن أعماله البارزة نظرية التقديرات، وتوزيعات المعاينة للعينات الصغيرة، وتحليل التباين وتصميم وتحليل التجارب. ومن العلماء الذين ساهموا كثيرا في نظرية التقديرات واختبارات الفروض كلاً من بيرسون Pearsson.E.s وكذلك نيمان Neyman – ويعد الثلاثي فيشر – بيرسون – نيمان مؤسس منهج الاستقراء الإحصائي والذي يعرف حاليا بالاتجاه الكلاسيكي. وهو يعتمد علي المعلومات المتاحة من العينة فقط.

وقد ظهر فى هذه الفترة اتجاه جديد يعرف بالاستقراء البيزيانى -Bayesianinfer ence وذلك بجهود كل من جفريز jeffreys ورافرى Ramsey وديفنتى Definetti وجودGood وسافج Savage ولندلى lindley وآخرون ويعتمد الاستقراء هنا على بيانات العينة بالإضافة إلى المعلومات المسبقة Prior. Information وشهدت هذه الفترة أيضا عملا مكثفا كان فيها الاهتمام منصبا على صنع القرارات، مما أدي إلى نشوء وظيفة حديثة للإحصاء تحت اسم نظرية القرارات الإحصائية -Decision the ory Statistical ويرجع ذلك إلى أعمال والد 1939) Wald) ونيومان Neuman.j ومورجنسترن Morgenstern.

وقد صاحب هذا التطور الكبير في النظريات الإحصائية بداية ظهور مجموعة من التخصصات المختلفة تهتم بمجالات وأهداف خاصة – وقد بلغ هذا التطور قدرا هائلا يكاد يظهرها وكأنها علوما مستقلة. ومن هذه التخصصات بحوث العمليات -Opera tions Research والإحصاء السكانى Demography ومراقبة الجودة Quality control والاقتصاد القياسى Econometrics ونظرا لاعتماد العلوم المختلفة على الرياضيات في فهم ظواهرها وقياسها وتفسيرها، فقد أفردت لها فروعاً خاصة تهتم بدراسة ظواهراها باستخدام الأساليب الإحصائية والرياضية ومنها على سبيل المثال الإحصاء الحيوى والاجتماع الرياضى والقياس الاجتماعي وعلم النفس الرياضى والقياس النفسى والقياس

التربوى والاقتصاد الرياضي والتاريخ الاقتصادي الجديد أو القياس التاريخى[1]

إن الأساليب الرياضية والإحصائية المستخدمة في مناهج البحث بصفة عامة تستخدم الآن في مجال العلوم الاجتماعية بنجاح. وقد أمكن عن طريقها التوصل إلى بعض الحقائق العلمية والنظريات، ولكنها لم ترق في هذا المضمار إلى ما وصلت إليه العلوم الطبيعية من نظريات علمية و قوانين.

وتصادف العلوم الاجتماعية صعوبات منهجية تحول دون تحقيق أهدافها في الوصول إلى ما وصلت إليه الأبحاث الطبيعية، ومن بين هذه الصعوبات.

- لا تخضع التفاعلات الاجتماعية لنظام آلى مرتب، ولا تسير وفق مبدأ الاطراد في تتابع الأحداث مما يسهل عملية الحصول على القوانين التى تحكم نظمها.
- صعوبة التوصل إلى قوانين التنبؤ الاجتماعي. وقد كان الاعتقاد السائد أن السلوك الاجتماعي والعلاقات الإنسانية التى تربط بين الأفراد في المجتمع إنما تخضع لنظم وقوالب يصب فيها الأفراد أعمالهم وأفكارهم ولا يكون الخروج عما ترسمه الطبيعة لهم من حدود وما تفرضه من التزامات.
- ليس لدى بعض العلوم الاجتماعية وحدات معينة تستخدم لقياس الظواهر موضوع الدراسة كما هو في العلوم الطبيعية التى تستخدم وحدات كمية لوصف ظواهرها والتعبير عنها بمعادلات رياضية والتنبؤ بها بتوافر شروط معينة.
- عدم استجابة البيئة الاجتماعية موضوع الدراسة للغايات التى يقصدها الباحث وعدم تمكن الباحث من السيطرة على كثير من العوامل التى تلعب دورا كبيرا في سير الحوادث وارتباط بعضها بالبعض الآخر.

(1) مصطفى زايد، مرجع سابق، ص 20.

والمزايا التى يجنيها الباحث من الطرق الإحصائية يمكن تلخيصها فيما يلي: -

- تساعد الباحث على إعطاء أوصاف على جانب كبير من الدقة العملية.
- فهدف العلم الوصول إلى أوصاف الظواهر و مميزاتها الطبيعية، وكلما توصل العلم إلى زيادة في دقة الوصف كلما كان هذا دليلا على التقدم العلمى ونجاح الأساليب العلمية. ودقة الوصف تحتاج دائما إلى اختبار مدى ثبات النتائج التى حصل عليها الباحث. فمجرد الوصول إلى نتائج دون التحقق من ثباتها لا يكفى عادة كأساس يعتمد علية في تفسير الحقائق وتحقيق الفروض.
- تساعد الإحصاء على تلخيص النتائج في شكل ملائم مفهوم فمجرد ذكر الدرجات لا يكفى للمقارنة بين الجنسين بل إن حساب متوسطى الدرجات قد سهل مهمة المقارنة كثيرا فالبيانات التى يجمعها الباحث لا تعطى صورة واضحة إلا إذا تم تلخيصها في معامل أو رقم أو شكل توضيحي كالرسوم البيانية.
- تساعد الباحث على استخلاص النتائج العامة من النتائج الجزئية.فمثل هذه النتائج لا يمكن استخلاصها إلا تبعا لقواعد إحصائية، كما يستطيع الباحث أن يحدد درجة احتمال صحة التعميم الذى يصل إليه.
- تمكن الباحث من التنبؤ بالنتائج التى يحتمل أن يحصل عليها في ظروف خاصة. فيما عدا الإحصاء يمكن للباحث أن يتنبأ بنتائج ما يجريه من اختبارات في وقت ما لقدرة أو قدرات خاصة لما ينتظر للأفراد الذين يختبرهم من نجاح في مهنة معينة أو نوع معين من التعليم.
- في كثير من البحوث يهدف الباحث إلى تحديد أثر عامل خاص دون غيرة من العوامل مما لا يتسنى تحقيقه عمليا. وهنا يستطيع أن يلجأ إلى الإحصاء فتعاونه على فصل عامل خاص من العوامل المحتمله وتحديد أثره على حده،كما تعينه على التخلص من أثر العوامل الأخرى التى لا يستطيع تفاديها في بحوثه والتى تؤثر دائما في نتائج كل بحث،كعامل الصدفة واختيار العينات.

- وقبل هذا كله تهدى الإحصاء الباحث عند تنظيم خطوات بحثه فهو يحتاج إليها في مرحلة تصميم البحث وتخطيطه،حتى يمكنه في النهاية أن يخرج من بحثه بالنتائج التى يسعى إلى تحقيقها، فهى تهديه إلى أضبط الوسائل التى تؤدى إلى التفكير الصحيح من حيث الإعداد أو الاستدلال والقياس أثناء خطوات البحث.

وإذا كان هو حال الإحصاء بالنسبة للبحوث العلمية بوجه عام فان حاجة البحوث الإنسانية أشد ما تكون إلى تطبيق هذه الوسائل. لذلك كانت البحوث النفسية والتربوية والاجتماعية من أصعب البحوث، وتحتاج إلى حرص زائد ومهارة فائقة من الباحث.

ويمكن تلخيص أسباب ذلك فيما يلي: -

أ) السلوك البشرى فى تغير دائم، ومدى تغيره من فترة لأخرى أوسع مما نظن، لدرجة تجعل من الصعوبة بمكان إعطاء تنبؤات علمية دقيقة عنه.

ب) السلوك البشرى كثيرا ما يخدع دارسة، ذلك لان حقيقته قد تختلف كثيرا عما يبدوا علية، فهو يحتاج إلى ضبط في البحث ودرجة كبيرة من الدقة الإحصائية.

ج) السلوك البشرى معقد تعقيدا كبيرا وتتدخل فيه عوامل قد تزيد أو تختلف عما يتوقعه الباحث.

د) البحوث الإنسانية يقوم بها إنسان. ذلك مما يسمح بتدخل العوامل الشخصية كثيرا في نواحي القياس والوصف بدرجة قد تكون كبيرة أو صغيرة حسب الطرق التى يستخدمها الباحث. وطرق الضبط الاحصائى خير وسيلة تعين الباحث على استبعاد هذه العوامل الشخصية.

إلا انه ينبغي أن يفهم من ذلك أن الإحصاء هو كل شيء في البحوث العلمية. فالإحصاء في يد من لا يجيد تطبيقها واستخدامها استخدام الخبير الفنى، لا تفيد كثيرا. فهى مرحلة تالية لاكتشاف المشكلة وتحديدها، وهى تتطلب عادة فروض علمية يتوقعها

الباحث بناءً على دراساته السابقة وملاحظاته العديدة، وهى تتطلب كذلك في آخر الأمر تفسيراً مبنياً على خبرة علمية وقدر وافى من المعلومات في الميدان الذى يجرى فيه البحث. وكلما كان الباحث مدركاً للأسس التى بنيت عليها الطرق الإحصائية التى يستخدمها، كلما سهل ذلك علية تطبيقها تطبيقا صحيحا، وتفسير النتائج تفسيراً مناسبا[1] ويتضح لنا من مفهوم الإحصاء أنة يمدنا بمجموعة من الأساليب والأدوات الفنية التي يستخدمها الباحث في كل خطوه من خطوات البحث ابتداء من المرحلة التمهيدية للبحث وما يتضمنه من عملية اختيار لعينة الدراسة وأسلوب جمع البيانات من الميدان ماراً بمرحلة تصنيف، وتلخيص، وعرض وتحليل تلك البيانات حتى مرحلة استخلاص نتائج الدراسة، ويرى البعض أن وظيفة الإحصاء يمكن أن تتلخص في نقطتين

الأولى: - تتمثل في تلخيص البيانات المتاحة وتقديمها في أبسط وأنسب صورة ممكنه. فالباحث عادة ما يجد نفسه أمام مجموعة كبيرة من البيانات الخام التي لا تفصح عن شئ على حين أنة مطالب باستخلاص حقائق علمية واضحة ومحددة من تلك البيانات سواء كانت بيانات مسوح اجتماعية شاملة. أو بالعينة أو بيانات تعدادات سكانية عندئذ يستطيع الباحث من خلال الإحصاء أن يغير من شكل البيانات بعد تصنيفها وتنظيمها وتلخيصها مستخدما فى ذلك الجانب الوصفى من الإحصاء حيث يمكنه أن يطبق هنا مجموعة من المقاييس الإحصائية التى لا تتعدى حد الوصف مثل مقاييس النزعة المركزية ومقاييس التشتت ومقاييس الارتباط والانحدار... الخ ومن ثم يتبين لدينا أن الوظيفة الإحصائية الأولى للإحصاء هى توصيف البيانات المتاحة والخروج منها بمجموعة من المؤشرات والمعـدلات الإحصائية.

الثانية: تتلخص في الاستدلال، ففى مجال البحوث الاجتماعية، عادة ما تستخـدم العينة sample لتمثل المجتمع الذى سحبت منه ويرجع استخدام العينات في البحوث الاجتماعية إلى عدة أسباب لعل أهمها توفير الوقت، والجهد، والإمكانيات التى تجعل من

(1) غريب محمد سيد أحمد، مرجع سابق، ص ص 14 – 18.

المتعذر أحيانا وربما من المستحيل أحيانا أخرى دراسة المجتمع ككل. والعينة ببساطة هى جزء أو قطاع من المجتمع تم اختيارها على أساس إحصائي لكى تمثل المجتمع الذى هى جزء منة وهنا يكون دور الإحصاء هو الوصول إلى تقديرات واستدلالات عن المجتمع ككل من خلال المعلومات المتوفرة عن العينة التى تم سحبها من هذا المجتمع، إذ إن جُل اهتمام الباحث ليس مجرد العينة المستخدمة في الدراسة بل المجتمع ككل، باختصار فان الجانب الاستدلالى من الإحصاء يهتم بتقدير معالم المجتمع Population Parameters فيما يتعلق بالظاهرة موضوع الدراسة مستخدما البيانات والمعلومات المتوفرة لدية عن العينة أو ما يسمى بـ Sample Statistics حول نفس الظاهرة في محاولة الوصول إلى تصميمات Generalizations عن مجتمع الدراسة.

هذا بالإضافة إلى اهتمام الإحصاء الاستدلالى باختبار الفروض العلمية. والإحصائية Hypotheses Teting للدراسة.

وإذا كانت تلك هى وظائف الإحصاء فى مجال العلوم الاجتماعية والتى يتضح منها بجلاء مدى ما تقدمة الإحصاء للباحث فهناك كلمة تحذير لابد أن يعيها كل من يفكر في استخدام الأساليب الإحصائية ألا وهى أن التطبيق غير الصحيح للأسلوب الاحصائى ربما يؤدى إلى نتائج غير صحيحة ومضللة كما أن استخدام الأساليب الإحصائية يجب ألا يكون غاية في حد ذاته بل انه وسيلة الهدف منها هو تبصير الباحث بما هو بصدد القيام به وتبسيط وتوضيح خطوات البحث العلمى[1].

وهكذا يتبين لنا مما سبق أن دراسة علم الإحصاء وان ثقلت على نفس بعض الأفراد، تعد ذات أهمية بالغة لأنها تزود الدارسين بالمهارات البحثية التى لم يعد أى فرض فى غنى عنها، ونحن نعيش عصر الثورة التكنولوجية وتهيمن على حياتنا لغة الأرقام[2].

(1) اعتماد علام، يسرى رسلان، مرجع سابق، ص ص 8 – 9.

(2) حسن محمد حسن، مرجع سابق، ص 20.

الفصل الثانى

المفاهيم الإحصائية

مقدمة

أولا: الإحصاء الوصفى والإحصاء الاستدلالي

ثانيا: البيانات

ثالثا: المتغيرات

رابعا: المقاييس الإحصائية

مقدمه

يزخر كل علم من العلوم بالعديد من المصطلحات والمفردات اللغوية الخاصة به والتى يعد الإلمام بها خطوة هامة على طريق الدراسة والفهم المتعمق لموضوعات ذلك العلم وعلم الإحصاء لا يختلف فى هذا الشأن عن غيره من العلوم فهو يتضمن عدد قليل من المصطلحات الأساسية التى نرى أن على الدارس أن يلم بتعريفاتها لكى يعي المقصود منها ويتسنى له معرفة كيفية التعامل معها عندما تعرض له فى دراساته وبحوثه ومن ثم يتفادى الخلط بين المصطلحات المختلفة عندما يحاول اختيار الأداة الإحصائية المناسبة لمعالجة البيانات التى قام بجمعها وتختلف الأساليب الإحصائية فيما بينها من حيث الهدف والتدرج من البساطة إلى التعقيد واختيار الأسلوب الملائم يتحدد وفقا لأهداف الباحث ونوعية البيانات المتاحة.

أولا: الإحصاء الوصفى والإحصاء الاستدلالي

(أ) الإحصاء الوصفى Descriptive statistics

ويهدف إلى إدماج وتلخيص البيانات الرقمية بغية تحويلها من مجرد كم من الأرقام إلى شكل أو صورة أخرى يمكن فهمها واستيعابها بمجرد النظر ومن أغلب الأساليب المستخدمة مقاييس النزعة المركزية، مقاييس التشتت ومقاييس الارتباط والانحدار ويتوقف استخدام أيا منها على نوعيه البيانات ومستوى القياس سواء أكان اسميا أو وصفيا، أو ترتيبا، أو فئويا، أو نسبة[1].

ويعتقد بعض الدارسين أن وظيفة الإحصاء تقتصر على معالجة مجموعة البيانات الوفيرة التى جمعها الباحث بقصد استخلاص عدد من الجداول الإحصائية وعرضها فى عدد من الأشكال والرسوم البيانية وذلك على نحو ما نشاهده فى إحصاءات السكان والاستهلاك والإنتاج وغيرها وقد يحسب المرء أن العمليات الإحصائية تدور فى جملتها حول إيجاد المتوسطات ودرجات التشتت فى البيانات التى يجمعها الباحثون ولكن فى الحقيقة أن ما ذكرناه لا يمثل سوى جانب واحد من جوانب الإحصاء وهو الجانب الوصفى ولهذا يطلق على العمليات الإحصائية التى تقوم بهذه الوظيفة مصطلح الإحصاء الوصفى وعلى هذا يستخدم الإحصاء الوصفى فى تنظيم وتلخيص ووصف معلومات خاصة بعينة من العينات فمن عينة محددة من العمال يمكن حساب متوسط الإنتاج الذى ينتجونه وحساب نسبة العمل بين أولئك العمال ومعدل الزيادة فى أجورهم وهذه المقاييس

(1) اعتماد علام، يسرى رسلان، أساسيات الإحصاء الإجتماعى، دار الثقافة للنشر والتوزيع، ص 10.

كلها وصفية بحتة لا تفيد فى حد ذاتها، فى الاستنتاج أو التنبؤ وإنما تصف الكيفية التى تتوزع بها البيانات التى تم الحصول عليها من العمال موضوع البحث[1].

وتعتبر وظيفة الوصف من الوظائف الأولية لعلم الإحصاء التى تستخدم فى تلمس حقائق الظواهر المختلفة (اجتماعية، اقتصادية، جغرافية.. الخ) وباستخدام أسلوب التحليل الاحصائى للبيانات أصبح من السهولة إمكان تحديد خصائص الظاهرة المدروسة حتى عن طريق الأشكال البيانية التى تمثل بيانات الظاهرية عملية تسهل وتبسط تحديد خصائص الظاهرة واتجاهاتها العامة.

والى جانب ذلك يعتمد الوصف فى الإحصاء على استخدام المقاييس والمؤشرات الإحصائية فى تقصى الحقائق وتحديد الخصائص العامة لتوزيع بيانات الظاهرة دون الوصول إلى نتائج أو استدلاله خاصة بالمجموعات الأساسية التى تنتمى إليها الظاهرة[2].

وعملية جمع البيانات تعد أقدم وظائف الإحصاء، وهى تتضمن عدد من الأنشطة يختلف مداها من مجرد بحث يقوم به فرد إلى فريق بحث من عدة مئات أو آلاف. وجمع البيانات يكون بعدد من الأساليب وحسب طبيعة البحث أو العمل، فقد يكون ذلك باستخدام المجموعات المكتبية أو عن طريق تصميم تجربة أو الملاحظة المنتظمة أو المعايشة أو عن طريق الاستبيان أو الاستبصار أو الأخبار بين الاختبارات ومهما يكن الأمر فإن جمع البيانات قد يتم إما بفحص كل وحدات المجتمع محل الدراسة أو بفحص جزئي (عينه).

إن عملية جمع البيانات ليست عملية منفصلة عن وظائف الإحصاء الأخرى فهناك صلة وثيقة – فالهدف واحد وهو الحصول على معلومات أو نتائج – وذلك يكون باستخدام مقاييس وأساليب وصف البيانات – وذلك بعد جمعها – وإذا كانت هذه البيانات خاصة بعينة أى بجزء من المجتمع فإن وصف المجتمع يتطلب استخدام أساليب الاستقراء.. وهذه

---

(1) حسن محمد حسن، أساليب الإحصاء وتطبيقاته، دار المعرفة الجامعية، 1992، ص ص 19 – 20.

(2) فتحى عبد العزيز أبو راضى، مبادئ الإحصاء الإجتماعى، دار المعرفة الجامعية، ص ص 2 – 3.

المقاييس والأساليب لها شروط ومتطلبات يجب مراعاتها وتوفيرها عند جمع البيانات وذلك باستخدام التصميم التجريبي المناسب أو تصميم استمارة استبيان مناسبة واختيار طريقة المعاينة المناسبة وحجم العينة المناسب ومراعاة توفير مستوى القياس المناسب للمتغيرات.. الخ كما أن البيانات التى يتم جمعها يجب أن تكون محل ثقة حتى تكون النتائج المستخلصة منها محل ثقة. أى يجب أن يتوافر فيها الصدق والثبات Validity and reliability أن تحديد ذلك واختياره يكون غالبا باستخدام الأساليب الإحصائية[1].

(ب) الإحصاء الاستدلالي Inferential Statistics

يستند هذا القسم من الأساليب الإحصائية إلى مجموعة من النظريات الإحصائية لعل أهمها نظرية الاحتمالات ونظرية العينات اللتان تمثلان حلقة الوصل بين الإحصاء الوصفى والاستدلالي. ويسعى هذا النوع من الأساليب الإحصائية إلى الوصول إلى تقديرات لمعالم وخصائص مجتمعات الدراسة من خلال ما هو متوفر من معلومات عن العينات المختارة. من تلك المجتمعات، فضلا عن اختبار الفروض الإحصائية عن مجتمع البحث على أساس البيانات المتاحة عن عينات الدراسة. ويطلق على هذا النوع من الأساليب أكثر من تسمية تؤدى جميعها إلى نفس المعنى فأحيانا يسمى بالإحصاء الاستدلالي، أو الاستنباطي Induc-tive أو التعميمي Generalizing حيث يهدف إلى الوصول إلى تعميمات عن مجمع الدراسة من خلال العينة المسحوبة من هذا المجتمع. ويشمل هذا النوع من الأساليب الإحصائية، الاحتمالات، العينات، اختبار الفروض، الاستدلال من خلال عينة واحدة أو أكثر وما يتضمنه ذلك من اختيارات مختلفة مثل كا2 chi2 اختبار جاما gamma، فاى phi ... الخ[2].

ويقصد بوظيفة الاستدلال اشتقاق النتائج من دراسة وفحص المقدمات والبيانات

(1) مصطفى زايد، الإحصاء ووصف البيانات، 1989، ص ص 26 - 27.

(2) اعتماد علام، يسرى رسلان، مرجع سابق، ص 10.

المتوافرة عن ظاهرة معينة. ولهذا يطلق علي عملية الإحصائية التي تستخدم والاستدلالي علي أساس المنطق الاستدلالي المبني علي نظرية الاحتمالات الرياضية فمن عينة محددة من أعمال أحد المصانع وباستخدام الأسلوب الإحصاء الاستدلالي يكون من الممكن التنبؤ بمعدلات الزيادة في الإنتاج ومقدار التغير في نسبة الغياب وفي هذه الحالة نجد أن الدقة في التنبؤ تعتمد علي عوامل كثيرة من أهمها ملائمة الأدوات الإحصائية المستخدمة وحجم العينة محل الدارسة والإجراءات الإحصائية اتخذت عند اختيارها[1].

وتعتبر وظيفة الاستدلال أو الاستقراء من الأهمية بمكان فى البحث العلمى فمثلا:

إذا كانت الظاهرة موضوع الدراسة والتحليل ممثلة للمجتمع الذى تنتمى إليه فانه يمكن الحصول على نتائج معنوية عن المجتمع بتحليل بيانات هذه الظاهرة وهو ما يعرف بالاستدلال ويعتمد هذا الأسلوب فى البحث على الشروط التى يجب توافرها حتى يكون هذا الاستدلال سليما – وبما أن الاستدلال لا يمكن أن يكون مؤكد فان لغة الاحتمال تستخدم عند عرض النتائج[2].

وتعتبر وظيفة الاستقراء لها أهمية كبيرة – فهى تمكن الباحث من الوصول إلى تعميمات عن المجتمع على أساس المعلومات المتاحة من عينة منه. وفى هذه الحالة فان أساليب ومقاييس الوصف يقتصر وصفها على ذلك الجزء (العينة) فقط من المجتمع – ومن هنا تأتى أهمية وظيفة الاستقراء – فهى تمكننا من وصف المجتمع (التعميم) باستخدام بيانات العينة.

إن القوانين فى العلوم الطبيعية والاجتماعية تجد برهانها عند الوقائع والحقائق الإحصائية ولذا يعد الاستقراء الاحصائى (Statistical Inference) أساسا لتطور المعرفة العلمية باعتباره البرهان لهذه القوانين. ووظيفة الاستقراء تحقق مطلبين أساسيين فى البحث:

(1) حسن محمد حسن، مرجع سابق، ص 36.

(2) فتحى عبد العزيز أبو راضى، مرجع سابق، ص 3.

الأول تقدير خواص المجتمع والثانى اختبارات الفروض حول هذه الخواص. ولا تقتصر هذه الوظيفة على مجرد الاستقراء بل تقدم لنا تقييما عن مدى دقة هذا الاستقراء وأكثر من ذلك فهى تمكننا من التحكم فى مستوى الدقة وذلك بعدة طرق منها استخدام الأسلوب المناسب للمعاينة والحجم المناسب للعينة. وباختصار فان هذه الوظيفة للإحصاء تمدنا بالاستقراء المنطقى وتختلف الأساليب المتبعة فى الاستقراء حسب طبيعة محل الاستقراء[1].

(1) مصطفى زايد، مرجع سابق، ص 32

ثانيا: البيانات Data

من الشائع فى مجال البحوث الاجتماعية توافر مجموعة من البيانات الإحصائية التى يحصل عليها الباحث باستخدام أدوات جمع بيانات مناسبة وعادة تتمثل تلك البيانات فى شكل أرقام تعتبر قياسا للمتغيرات تحت الدراسة ولما كانت تلك الأرقام تفتقر إلى الترتيب والتصنيف يطلق عليها البيانات الأولية أو البيانات الخام Raw Data.

وتعرف البيانات الإحصائية أنها كمية من المعلومات على هيئة أرقام وان تلك الأرقام إما أن تكون صحيحة Integers مثل 10، 20، 30 وهكذا أو تكون أرقاما عشرية أو حقيقية Real Numbers مثل 8.5، 10.25، 1505 وهكذا: ويتوقف حجم البيانات الخام على حجم المجتمع الأصلى فكلما ازداد حجم هذا المجتمع يتوقع مزيدا من الأرقام غير المرئية والتى يصعب مع كثرتها وعدم تصنيفها تفهم أو قياس متغير أو أكثر تحت الدراسة ومن ثم كان من الضرورى أن يقوم الباحث بتصنيف وتبويب تلك البيانات بالشكل أو بالأسلوب الذى يخدم جيدا هدف الباحث من دراسة المتغيرات أو استنباط نوعية العلاقات أو المعلومات الهامة التى تتعلق بتلك المتغيرات[1].

ويقصد بتعبير البيانات « أى كمية من المعلومات فى صورة رقمية والصورة الرقمية للبيانات تبدو إما على شكل أرقام صحيحة مثل 10، 112، 464. أو على شكل أرقام حقيقية مثل 20.4، 61.8، 182.1 أى أنها الأرقام التى تحتوى على علامة عشرية. وتعتبر المعلومات الرقمية (البيانات) المادة الخام لأسلوب العمل الاحصائى كما أنها تلعب دورا كبيرا فى تطبيق الأساليب الإحصائية[2].

(1) اعتماد علام، يسرى رسلان، مرجع سابق، ص34

(2) فتحى عبد العزيز أبو راضى، مرجع سابق، ص 6

وتسمى البيانات المتاحة – المنشورة أو التى تم جمعها – تسمى بيانات خام أو أولية – ذلك أنها تكون غير مجهزة فهى لا تفصح إلا عن القليل من المعلومات. كما أنه يستحيل استخلاص المعلومات منها. وفى سبيل ذلك نستعين بأساليب ومقاييس وصف البيانات. وهذه الأساليب كثيرة ومتنوعة فهى تختلف حسب عوامل أهمها عدد المتغيرات ومستوى قياسها[1].

ولعل ابسط الطرق الإحصائية لتنظيم وتلخيص البيانات طريقة التوزيع التكرارى Frequency Distribution، أو بمعنى ضمنى من التوزيع التكرارى يمكن استخدام وسيلة أو أكثر من الوسائل الثلاث التالية والتى يمكن أن يتحول التوزيع إليها أو إلى أى منها.

أ) استخدام الجداول الإحصائية Statistical Tables فى عملية تصنيف وتبويب البيانات الخام.

ب) استخدام التمثيل البياني والخرائط فى عرض البيانات الإحصائية (تحويل التوزيع التكرارى إلى منحنيات تكرارية).

ج) استخدام مقياس أو أكثر من المقاييس الإحصائية مثل المتوسط الخام Mean الانحراف المعياري Standard Deviation ومعامل الارتباط Correlation Coefficient فى تلخيص البيانات الإحصائية فى صورة رقم أو نسبة مئوية ونرى أهمية الوقوف على نوعية البيانات الإحصائية من منظور مستويات القياس الاحصائى نظرا لأهمية تلك البيانات الإحصائية وفقا لمستويات القياس الاحصائى يرجع إلى أن المتغيرات التى تقاس كميا تنقسم من قيمتها العددية إلى المتغير المتصل والمتغير المتقطع[2].

---

(1) مصطفى زايد، مرجع سابق، ص 30

(2) اعتماد علام، يسرى رسلان، مرجع سابق، ص ص 34 – 35

ثالثا: المتغيرات Variables

تشير كلمة المتغيرات إلى الخصائص التى تشترك فيها أفراد المجتمع الاحصائى ولكنها تختلف من فرد إلى فرد آخر فالعمر، درجة الذكاء، وطول القامة، واللياقة البدنية والقدرة على القراءة، والدخول التى يحصل عليها الأفراد أمثلة للمتغيرات وتتميز هذه المتغيرات بأنها قابلة للقياس الكمى وبإمكانية تحديد قيمة معينة لها.

ويمكن القول بان المتغيرات مفهوم له معنى امبريقى ويعبر عنه بقيم مختلفة وتعبير النوع، سنوات التعليم والعمر، والدخل السنوى من المتغيرات الشائعة التى تستخدم فى البحوث الاجتماعية لارتباطها بالخصائص الأساسية للمبحوثين، ولأهميتها فى تحديد مكانتهم الاجتماعية والاقتصادية وانتماءاتهم الطبقية[1].

والمتغيرات عبارة عن ظاهرات أو صفات تختلف قيمها باختلاف الحالات. ومن أمثلتها: درجة الحرارة فى مناطق مختلفة أو فى فترات مختلفة لمكان واحد، كميات الإنتاج الزراعى أو الصناعى[2].

ويمكن القول بان المتغير هو أى ظاهرة أو حدث أو خاصية تأخذ فيها قيما تتغير من ظرف لآخر. والمتغير هو الوحدة الأساسية للتحليل الاحصائى ويمكن تعريفة بأنه مجموعة من العناصر أو التقسيمات غير المتداخلة. وهذه المجموعة من التقسيمات تكون مقياس Scale. وتنقسم المتغيرات إلى مستمرة وغير مستمرة (متقطعة). المتغير المستمر هو ذلك الذى يأخذ قيما لأى درجة من الدقة – مثل الطول – الوزن – درجة الحرارة أما المتغير

(1) حسن محمد حسن، مبادئ الإحصاء الاجتماعي، دار المعرفة الجامعية، ص ص 36 – 37

(2) فتحى عبد العزيز أبو راضى، مرجع سابق، ص 7

غير المستمر فهو الذى يأخذ قيما معينة فقط – مثل عدد الأولاد فى الأسرة عدد الطلاب فى الفصل. وهناك تقسيم آخر للمتغيرات، حيث تنقسم إلى متغيرات مستقلة ومتغيرات تابعة. فعندما نبحث فى الأثر الذى يحدثه متغير (س) فى آخر (ص) كأثر التدريب على الإنتاجية نقول أن (س) متغير مستقل و (ص) متغير تابع[1].

وتنقسم المتغيرات من قيمها العددية إلى قسمين هما المتغيرات المتصلة Continuous Variables وهى المتغيرات التى يمكن أن تأخذ أى قيمة على المقياس المستخدم فمثلا إذا ارتفعت درجة الحرارة من 520 درجة مئوية إلى 530 درجة مئوية خلال الترمومتر الزئبقي فمعنى ذلك أن الزئبق يكون قد مر بكل القيم الواقعة بين هاتين الدرجتين، كذلك الحال فى مقياس سرعة السيارة. فإذا زادت السرعة من 30 كيلوا متر / ساعة إلى 60 كيلوا متر / ساعة فان المؤشر فى المقياس يكون قد مر على كل القيم المحصورة بين هذين الرقمين وبالمثل أيضا الأطوال. وذلك لان طول الشخص قد يكون 168 سم أو 168.1 أو اى قيمة مهما كانت كسرية، واصغر من المليمتر إذا كان المقياس يسمح بذلك.

والنوع الآخر من المتغيرات يطلق علية المتغيرات الغير متصلة أو الوثابة Discrete Variables وهى التى تختلف قيمها من مرحلة إلى أخرى بدون أن تكون منتظمة كما أن قيمها لا تأخذ إلا أعداد صحيحة Integers فعدد الرحلات التى يقوم بها الأشخاص وكمية مياه الفيضان فى الأودية الصحراوية وعدد السيارات المارة فى احد الشوارع وعدد الفصول بالمدارس وعدد الحجرات بالمنازل وحجم الأسرة 000 الخ كلها متغيرات وثابة (غير متصلة) يحصل عليها فى الغالب بالعد[2].

والمتغيرات التى تقاس كميا تنقسم من حيث قيمتها العددية إلى نوعين هامين لا ثالث لهما:

---

(1) مصطفى زايد، مرجع سابق، ص ص 23 – 24

(2) فتحى عبد العزيز أبو راضى، مرجع سابق، ص ص 7 – 8

1 – المتغير المتصل Continuous Variable.

لما كان التعريف العام للمتغير Variable هو ظاهرة أو صفات تختلف قيمها باختلاف الحالات فان المتغير يكون متصلا عندما يأخذ أى قيمة متدرجة على المقياس المستخدم. مثال ذلك قياس درجات الحرارة باستخدام الترمومتر فالمتغير يأخذ أى قيمة بين رقمين صحيحين . بمعنى أن المتغير يمكن أن يأخذ أى قيمة بين 36 درجة، 37 درجة (36.1، 36.2 000 الخ).

2 – المتغير المتقطع Discrete Variable

عندما يأخذ المتغير قيما محددة يطلق عليه متغيرا متقطعاً أو بمعنى آخر، المتغير المتقطع هو الذى يحتوى مداه على عدد محدود من القيم أو يحتوى عدد لانهائي من القيم ولكن لكل منها قيمة محددة يمكن عدها أو ترتيبها فى نهاية الأمر تعدد الأولاد أو الأفراد فى الأسرة لابد أن يكون أعدادا صحيحة غير حقيقة مثل 1، 2، 3، 4 00 وهكذا ومن أمثال المتغيرات المتقطعة، النوع، الحالة الزواجية Martial Status . عدد أيام الإنتاج فى احد المصانع، عدد حوادث السيارات وهكذا[1].

كما يمكن تصنيف المتغيرات إلى عدد من التصنيفات بحسب الغاية من كل تصنيف وذلك على النحو التالى: –

1 – المتغيرات الكمية والمتغيرات الكيفية:

يمكن تصنيف المتغيرات من حيث طريقة التعبير عنها إلى فئتين هما: المتغيرات الكمية Quantitative Variables وهى التى يمكن أن نصفها عدديا بأنها اكبر من أو أقل من قيمة معينة ويعتبر العمر وعدد سنوات التعليم أمثلة لهذه المتغيرات. والفئة الثانية من المتغيرات هى المتغيرات الكيفية Qualitative Variables وهى التى تصف الأشياء بصفاتها مثل

(1) اعتماد علام، يسرى رسلان، مرجع سابق، ص 35

متغير النوع الذى ينقسم إلى قسمين: ذكور وإناث. والحالة العملية للفرد حيث تكون إما مزارع أو عامل غير ماهر، أو عامل ماهر أو موظف أو تاجر وما إلى ذلك من صفات، وهذه المتغيرات الكيفية يتعذر معالجتها إحصائياً ما لم يميزها عن بعضها بعضاً باستخدام الأرقام فنرمز لمتغير الإناث برقم 1 و لمتغير الذكور برقم 2 أو العكس، والرقم فى هذه الحالة لا يعنى أكثر من أنه أداه للتميز بين المتغيرات الكيفية لتسهيل تفريغ البيانات التى جمعت عنها من ميدان الدراسة تمهيداً لمعالجتها إحصائياً ولا تكون لها قيمة عددية فى حد ذاته.

2 – المتغيرات التابعة والمستقلة والضابطة:

ويمكن تصنيف المتغيرات تصنيفاً أخر بحسب دورها فى حدوث الظاهرة محل الدراسة وذلك إلى:

(أ) متغيرات تابعة Dependent Variables

وهى تلك المتغيرات التى نحاول تفسيرها ومعرفة أسباب حدوثها وتحديد مدى إمكان التنبؤ بها.

(ب) متغيرات مستقلة Independent Variables

وهى التى لعبت دوراً مباشر فى حدوث المتغيرات التابعة ونستخدمها فى تأييد تفسيرنا وفهمنا لما طرأ على هذه المتغيرات من تغيير، وفى التنبؤ بالحالة التى ستؤول إليها بعد ذلك.

(ج) متغيرات وسيطة Intermediate Variables

وهى تلك المتغيرات التى يمر من خلالها تأثير المتغيرات المستقلة إلى المتغيرات التابعة والمتغيرات الوسيطة بالغة الأهمية فى تفسير حدوث الظواهر الاجتماعية إذ قد يغفل عنها الباحثون أو قد ينظرون إليها على أنها متغيرات مستقلة لارتباطها المباشر بالمتغيرات التابعة فإذا نظرنا إلى تفسير ظاهرة الانتحار اللامعيارى التى درسها دوركايم، على سبيل المثال سنجد أن بعض الأفراد ينظرون إلى حالة فقدان المعايير التى تؤدى إلى الانتحار على أنها

المتغير المستقل والانتحار هو المتغير التابع ولكن فريقاً آخر من الباحثين الذين ينظرون إلى الظاهرة بطريقة أكثر تفصيلاً، ويرون أن المجتمع يمر بتغيرات اقتصادية و اجتماعية عاصفة وقوية وهى التى تمثل المتغير المستقل وتكون النتيجة المترتبة على تلك التغيرات انهيار الثقة فى القيم الراسخة والمبجلة لدى الأفراد فتنتشر حالة اللامعيارية وهى تمثل هنا المتغير الوسيط ثم ينتهي الأمر بالانتحار الذى يمثل المتغير التابع. وإذا قارنا بين الطريقتين السابقتين فى تفسير ظاهرة الانتحار نجد أن حالة اللامعيارية كانت متغيراً مستقلاً فى التفسير الأول ثم اعتبرت متغيراً وسيطاً ضابطاً فى التفسير الثانى.

3 – المتغيرات غير المستمرة (الوثابة)، والمستمرة (المتصلة) Discrete and continuous variables

ذكرنا أن مهمة الباحث هى جمع البيانات عن متغيرات معينة مثل متغير النوع بأن يعرف كم عدد المبحوثين من الذكور وكم عددهم من الإناث، وعن متغير سعة الوحدة السكنية بأن يحدد عدد الغرف التى يسكن بها كل مبحوث.

وبالنظر إلى المتغيرات السابقة نجد أنها تضم عدداً من المتغيرات غير المستمرة والتى يمكن التعبير عنها بقيم عددية غير قابلة للتجزئة حيث يرمز الباحث للذكور برقم (1) وللإناث برقم (2)، ولا توجد قيمة وسط بينهما وكذلك الحال بالنسبة لسعة الوحدة السكنية، فالشقة إما أن تكون غرفة واحدة أو غرفتين أو ثلاث أو أكثر وليس هناك جزء من غرفة. والبيانات التى يتم جمعها عن المتغيرات غير المستمرة تكون بيانات غير مستمرة أيضاً أى أنها غير قابلة للتجزئة ولا نجد لها كسور. فلا يستطيع الباحث أن يدعى أن العينة تتكون من عشرة ذكور ونصف أو أن الشقة تتكون من ثلاث غرف وربع. ويطلق على البيانات الكمية التى يتم جمعها عن المتغيرات غير المستمرة القيم المفردة حيث لا يمكن تبويبها أو تقسيمها إلى فئات متصلة.

وقد يهتم الباحث أيضاً بجمع بيانات عن دخل كل مبحوث فى فترة معينة. والدخل

يعد من المتغيرات المستمرة التى يمكن أن تأخذ أى قيمة ما بين نقطتين ثابتتين على مقياس معين. وإلى جانب الدخل هناك متغيرات أخرى مثل العمر والطول والوزن تعد أيضاً من المتغيرات المستمرة، إذ يمكن تقسيم متغير كالدخل إلى أى عدد نشأ من الفئات وكذلك متغير العمر فيمكن القول أن هناك شخصاً يحصل على دخل أسبوعى قدره خمسون جنيهاً وآخر يحصل على تسعة وأربعون جنيها ونصف ... وهكذا والبيانات التى يتم جمعها عن المتغيرات المستمرة تكون بيانات مستمرة أيضاً أى أنها قابلة للتجزئة وبها كسور أو قيم غير صحيحة.

ولذلك فإن هذا النوع من البيانات الكمية يكون ضخماً للغاية عندما يجمعه الباحث من ميدان البحث. فإذا سأل مائة فرد عن دخلهم الأسبوعى فأنه من المتوقع أن يحصل على مائة إجابة تمثل مائة قيمة مختلفة عن بعضها البعض. ولذلك عادة ما يتم تفريغ هذه البيانات فى صورة فئات لكل منها طول معين بحيث تحتوى كل فئة على عدد من القيم المتقاربة لتسهيل عرض البيانات ومعالجتها إحصائياً، وهذا النوع من البيانات نطلق عليه البيانات أو القيم المبوبة.

والواقع أن التمييز بين المتغيرات غير المستمرة والمستمرة رغم أهميته إلا أنه فى بعض الأحيان نظراً لعدم وجود أداة قياس مضبوطة نجد أن متغيرات كثيرة مستمرة يكون من الضرورى تحديد قيم عددية إجمالية لها، ومن ذلك مثلاً مقياس الذكاء فهو من الناحية النظرية يعد متغيراً مستمراً ولكن من الناحية العملية نجد أن الاختبارات التى تستخدم فى قياسه تعطى نتيجة إجمالية وقيمة غير مستمرة[1].

(1) حسن محمد حسن، مرجع سابق، ص ص 37 – 40

رابعا: المقاييس الإحصائية

يقصد بالقياس - كمفهوم واسع - انه عملية تعبير عن الخصائص والملاحظات بشكل كمى ووفقا لقاعدة محدودة. وعندما نستخدم المقياس والملاحظات بشكل كمى ووفقا لقاعدة محددة. أو بمفهومه وفق الأبعاد الخاصة الملائمة لكل فرع من فروع المعرفة، فإننا لا نجد غضاضة فى اختيار نسق من المعادلات الرياضية التى تتفق مع تلك الخاصية أو الخصائص قيد البحث - وعامة يمكن القول أن ما تحظى به فروع العلم المختلفة من رياضيات واقتصاد وغيرها من فروع العلوم الاجتماعية من نماذج متعددة ومتباينة تعتمد فى بنيتها الأساسية على المقاييس.

وإن كان هناك اختلاف كبير فى درجة الصعوبة عند التطبيق إذا قورنت النماذج المستخدمة فى العلوم الاجتماعية بغيرها من فروع العلوم الأخرى ففى علم الاجتماع وعلم النفس الاجتماعي كمثال تتصف المتغيرات بالتباين والتعدد بشكل يصعب معه أن نختار رياضيا مناسبا يخدم أهداف البحث الامبريقى لان النفس البشرية (والفرد عامة) - يتصف بالتعقيد واختلاف مستويات العلاقة بينة وبين المحيطين به من أفراد أو بيئات

ولعل ابسط أمثلة القياس نجدها فى الاختبارات التى يتقدم بها الطالب فى مختلف مراحل حياته الدراسية. حيث ترتبط الدرجة التى يحصل عليها فى اختبار على مدى معرفته بالمادة التى يدرسها خلال فترة دراسية معينة وكلما كانت درجة الطالب التى حصل عليها مثلا فى مادة الكيمياء عالية دل ذلك على معرفة أكثر أو تحصيل اكبر لدى الطالب من هذه المادة. ومن هذا المثال البسيط نجد أن خاصية التحصيل تعبر عنها الدرجة Score التى حصل عليها الطالب من الاختبار.

وتعتبر المقاييس التى تقيس المتغير التابعDependent Variable واحدة من أكثر المقاييس أهمية عند إيجاد الطرق الإحصائية الملائمة التى تستخدم فى تحليل بيانات دراسة أمبريقية معينة. أيضا توجد بعض المقاييس التي يمكن استخدامها في قياس ظاهرة معينة بدقة عالية أو متناهية مثال ذلك المقاييس التي تستخدم في قياس الأطوال والأوزان من جهة أخري توجد بعض المقاييس التي تفتقر إلي الدقة العالية وإن كانت تحقق قدرا من الدالة فيها علي سبيل المثال مقاييس مستويات القلق النفسي عند الأفراد[1] ويعتمد القياس في التحليل الإحصائي علي القيم العددية التي تستخدم بطرق مختلفة لتحقيق عدة أهداف: -

1 - تستخدم القيم العددية لترقيم المتغيرات (إجابات الأسئلة) التي يختار من بينها المبحوث في الاستبيان المكتوب.

2 - وتستخدم القيم العددية في ترتيب مجموعة من المتغيرات فيكون المتغير رقم (1) أعلي من المتغير رقم (2) عندما يكون الترتيب تنازلي للقيم ويكون المتغير رقم (1) أدني من المتغير رقم (2) عندما يكون الترتيب تصاعدي للقيم بعبارة أخري . تفاوت أهمية القيم بحسب ما إذا كان الترتيب تصاعديا أو تنازليا.

جـ - تستخدم القيم العددية أيضا في تحديد المسافة بين الفئات المختلفة من المتغيرات لذلك يجب علي الباحث أن يفهم الكيفية التي تستخدم بها الإعداد في وضع المقاييس الإحصائية[2].

ولغرض استخدام المقاييس والأساليب الإحصائية فإنه يجب تحديد مستوي القياس للبيانات أو المتغيرات ولذلك يتم تقسيم مستويات القياس إلي أربعة أنواع هي مستوي القياس الاسمي والترتيبي والفتري والنسبي وهذه المقاييس تختلف من حيث كمية المعلومات التي تحتويها وبالتالي تختلف العمليات الحسابية والإحصائية التي يمكن إجراءها[3].

(1) اعتماد علام، يسرى رسلان، مرجع سابق، ص ص 35 - 36

(2) حسن محمد حسن، مبادئ الإحصاء الاجتماعي، مرجع سابق، ص 40

(3) مصطفى زايد، مرجع سابق، ص 24

1 – المقاييس الاسمية والوصفية nominal measures هذا النوع من المقاييس يستخدم المتغيرات التي تستخدم في تصنيف مفردات عينة البحث وذلك بإعطائها قيما عددية والقيمة العددية في هذه الحالة ليس لها دلالة سوي تعريف المتغيرات وتمييزها ويستعين بعض الباحثين بالرموز بدلا من الأرقام في عملية استخدام المتغيرات في تصنيف بعض مفردات عينة البحث ولكن استخدام الرمز لن يفيد كثيرا في حالة تفريغ البيانات بواسطة الحاسب الآلي ومن أمثلة المتغيرات التي تشكل منها المقاييس الوصفية التي تستخدم في تصنيف المبحوثين متغير النوع إذا يعطي الباحث رقم (1) للإناث ورقم (2) للذكور أو يصف المبحوثين حسب متغير الدين إلي (1) مسلم (2) مسيحي (3) يهودي – والأرقام هنا لا تعني أولوية أو أفضلية متغير علي أخر كما أنها لا تحتمل أي قيمة. والواقع أن أرقام السيارات وأرقام المنازل هي أبرز مثال لاستخدام القيم العددية في تصنيف الأشياء فالمنزل رقم (1) ليس يعني أنه أفضل من المنزل(100) أو العكس وإنما الرقم يكون استخدامه بغرض التعرف علي المنزل وتميزه عن المنازل الأخرى[1] ويعد أقل مستوي للقياس . وهو مجرد تقسيم أو تصنيف الأشياء بالاسم فقط ودون تداخل مثال ذلك تقسيم الأشخاص حسب الجنس (ذكور – إناث) وحسب الجنسية (مصري – سعودي – عراقي..........) وتقسيم الجرائم إلي (قتل – خطف – سرقة) وتقسيم الكتب والمراجع بالمكتبة حسب الموضوع (المعارف العامة – الفلسفة – الديانات – العلوم الاجتماعية) وتشمل قياسات خصائص الظاهرة موضوع الدراسة في هذا النوع علي قياسات[2] ثنائية أو ثلاثية ولنضرب مثالا علي ذلك فعند تسجيل حالة التعليم لدي الأشخاص: تعليم متوسط أن تعليم عالي يعطي الشخص من النوع الثاني الرقم (2) وإذا كانت الحالة التعليمية يعطي الرقم (صفر) . وإذا كانت الدراسة تتعلق بانتماء الأشخاص إلي مناطق ريفية أو حضرية فإننا في هذه الحالة نعطي للشخص الريفي الرقم (1) وللشخص الحضري الرقم (2) ويطلق علي المتغيرات التي تقاس بها البيانات

(1) حسن محمد حسن، مبادئ الإحصاء الاجتماعي، مرجع سابق، ص 41

(2) مصطفى زايد، مرجع سابق، ص 25

الاسمية المتغيرات دمي dummy variables كما أنها في أحيان أخري تسمي بالبيانات التصنيفية لأنها تصنف المتغيرات علي أساس خصائصها[1].

ويعتبر التصنيف أبسط العمليات الأساسية في أي فرع من فروع العلم فالتصنيف هو تجميع للمفردات أو العناصر أو المعلومات المتشابهة إلي حد كبير المتماثلة في خصائصها مع بعضها في مجموعة أو مصنف category وذلك بهدف المقارنة بين المجوعات المختلفة علي أساس الخواص مثال ذلك إذا قمنا بتصنيف عدد من الأفراد إلي مجموعات وفق خاصية العقيدة religion (مسلم – مسيحي – يهودي) وقد تقوم أيضا بعمل تصنيف أخر للنزعات السياسية للفئات الدينية الثلاث وهكذا ولابد من استخدام التصنيف كعملية أساسية تعتمد عليها المقاييس الأعلى كأساس لها أيضا في العلوم الاجتماعية من ذلك لا نبالغ بالقول إن التصنيف يعتبر المستوي الأول في القياس وفي المثال السابق نجد أننا لم نهتم بالتمييز بين الفئات الدينية الثلاث علي أساس الأهمية مثلا فلم نقل أن المسلم أهم من المسيحي أو أن المسيحي أهم من اليهودي فقط ينصب المقياس علي تصنيف وفق الديانة وتمثل الخاصية الأولي للمقياس التصنيفي والتي يمكن أن نحددها في عدم اتصاف المقياس بالترتيب المنطقي من ذلك نلاحظ عدم وجود أي تدخل علي أساس الديانة فالمجموعة كاملة تضم أفراد متماثلين في نوع الديانة ومن ثم لا تتكرر الظاهرة أو المفردة في أكثر من مجموعة وهذه ميزة ثانية وهامة يتصف بها المقياس التصنيفي والخاصية الثالثة التي تتصف بها المقاييس التصنيفية نجدها في مجال العلاقات بين المفردات أو المقادير في العلوم الرياضية علي سبيل المثال يتصف المقياس بخاصية الانتقالية transitivity ويقصد بها أنه إذا كانت هناك علاقة معينة بين متغيرين من أ.ب بحيث أنها تتحقق من (أ) (ب) فإن من الضروري أن تتحقق أيضا من المتغير (ب) نحو المتغير (أ)[2].

(1) فتحى عبد العزيز أبو راضى، مرجع سابق، ص 8

(2) اعتماد علام، يسرى رسلان، مرجع سابق، ص 37

2 – المقاييس الترتيبية ordinal measures وهذه المقاييس لا تستخدم فقط لتصنيف المتغيرات وإنما لتعكس أيضا ترتيب تلك المتغيرات بعبارة أخري يستخدم هذا المقياس في ترتيب الأفراد أو الأشياء من الأعلى أو العكس وذلك وفقا لخصائص معينة يتميز بها المراد ترتيبه فالمكانة الاجتماعية – الاقتصادية والتي تقاس بمتغيرات الدخل والمهنة والتعليم يتم ترتيبها حسب فئات معينة تبدأ تنازليا من الطبقة العليا الطبقة عليا الوسطي – الطبقة الوسطي الطبقة وسطي الدنيا – والطبقة الدنيا – ما دون الطبقة under class فإذا أعطينا أرقاما لهذا الترتيب الطبقي فإن رقم (1) يكون له معني يفيد الرقمي إذا ما قورن برقم (4) وهكذا ويستخدم هذا المقياس أيضا في وصف المتصلات continuums مثل المتصل الريفي – الحضري الذي يكون بدايته رقم 1 – الريف 2 – الأطراف الحضرية 3 – الحضر 4 – الضواحي فرقم (1) هنا يشير إلي بداية المتصل ورقم (2) يشير إلي مرحلة أخري منه وهكذا الحال بالنسبة لباقي المتصل[1].

وهذا القياس أعلي مستوي من المقياس الاسمي حيث يتم التقسيم علي أساس الرتبة أو الأهمية النسبية مثال ذلك درجات الطلاب علي أساس ممتاز – جيد جدا – جيد – مقبول – ضعيف أو توزيع السكان حسب الحالة التعليمية: أمي – ابتدائي – ثانوي – جامعي – ماجستير – دكتوراه وفي هذا القياس يمكن ترتيب القيم وإجـراء المقارنات حيث يمكن القول أن الحاصل علي تقدير جيد مستوي تحصيله أفضل من الحاصل علي تقدير مقبول مثل هذا الترتيب والمقارنة لا نستطيع القيام بها في المقياس الاسمي حيث أن هذا المقياس لا يمكنه تحديد مقدار الفروق بين القيم[2] وتعرف القياسات الترتيبية بالبيانات المرتبة في فئات أو حسب خصائصها عن طرق إعطاء القيم الأصلية للمتغيرات رتبا أو أرقام تدريجية أو تنازلية[3].

(1) حسن محمد حسن، مبادئ الإحصاء الإجتماعى، مرجع سابق، ص ص 41 – 42.

(2) مصطفى زايد، مرجع سابق، ص ص 24 – 25.

(3) فتحى عبد العزيز أبو راضى، مرجع سابق، ص 9.

وفضلا عن تصنيف الأفراد إلي ثلاث مذاهب دينية يمكن أن ترتب تلك المجموعات الثلاثة وفقا لأهميتها أو لما تمتلكه كل منها من خاصية أو سمات معينة مشتركة وغير مشتركة وقد نجد مثالا أقرب للفهم في الرياضيات عندما نميز بين المقدارين (أ)، (ب) فنقول أن (أ) > (ب) ونأخذ الشكل الرياضي التالي أ > ب وقد يكون أ < ب ولكن مقدار الفرق في القيمة الدالة علي التمييز بين أ . ب ليس من خصائص المقياس الترتيبي ومن ثم فإن المقياس الترتيبي هو مستوي أعلي من المقياس التصنيفي في قياس الظواهر أو الخواص وتعتبر خاصية التمييز باستخدام علامات(<) أو (>) الخاصية الثانية إذا أخذنا في الاعتبار الخاصية التصنيف وفق الترتيب وفي العلوم الاجتماعية نجد مثالا لخاصية الترتيب دون الالتزام بالفروق عندما نصنف الأسر وفقا للمكانة الاجتماعية الاقتصادية socio eco-nomic status طبقة عليا . متوسط عليا upper middle . متوسط دنيا lower middle وأيضا إلي طبقة دنيا lower class وحقيقة الترتيب هنا هما الرتبة العليا والرتبة الدنيا فقط والخاصية الثالثة لو تخيلنا ترتيبا للأفراد علي متصل continue شريطة ألا يحتل فردان منهما مكانا واحدا أو يتواجدان في نقطة واحدة علي هذا المتصل وذلك مع فرض وجود علاقة أو روابط بين هؤلاء الأفراد علي المتصل ومن ثم يتم جيعهم عشوائيا دون دراية كافية في مجموعة وتكرار ذلك وفق ترتيب لخاصية معينة بحيث يمكن لنا فقط أن نقول أن المجموعة كذا من الأفراد تمثل أعلي التكرارات قياسا بباقي المجموعات أو نقول أن المجموعة كذا تمثل أعلي النقاط نسبيا هذا ويجدر الإشارة أن جميع المفردات دون تكرار ظهور المفردة في أكثر من مجموعة تمثل خاصية يتشابه فيها المقياس الترتيبي مع المقياس التصنيفي والخاصية الرابعة فهي الانتقالية فلو فرضنا قريبا أن أ > ب وأن ب > جـ وهذه خاصية أخري يتشابه فيها هذا المقياس مع المقياس التصنيفي ولكن من المنظور الترتيبي ويجب التنويه إلي ضرورة ملاحظة أن المستوي الترتيبي للقياس لا يهتم بالفروق – كما قلنا – بين العناصر أو الخواص ومن ثم لا نستطيع أن نستخدم مع هذا المقياس التصنيفي ولتوضيح ذلك

فالعمليات الحسابية كالطرح والقسمة والضرب والجمع لا يمكن استخدامها أيضا مع المقياس التصنيفي وبافتراضنا أن هناك أربع نقاط متصلة ويرمز لها بالأحرف (أ.ب.ج.د) ويفارق مسافات معينة تقع النقطتان ب.ج بين النقطتين (أ) .(د) في الشكل التالي متصل

أ ب جـ د

فباستخدام المقياس الترتيبي يمكن كتابة العلاقة التالية (اتجاهيا).

أد = أب+ ب جـ+ جـ د ولكن لا يمكن إطلاقا معرفة أطوال المسافات الأربعة المبينة في العلاقة السابقة مثال ذلك الترتيب المستخدم في مقاييس الاتجاهات الذي يبدأ بالموافقة بشدة وينتهي بعدم الموافقة بالمرة[1].

3 – مقاييس الفئات Interval measures

يشير مقياس الفئات إلي تبويب البيانات وتقسيمها إلي رتب معينة تبدأ من أدني الفئات إلي أعلي الفئات، وبالإضافة إلي ذلك فهو يحدد المسافة بين تلك الرتب وتستخدم مقاييس الفئات في تلخيص القيم المتقاربة لتكون فئة واحدة، ويعتبر الدخل، والتعليم ودرجات الحرارة والعمر أمثلة علي المتغيرات التي تستخدم في تبويب بياناتها مقاييس الفئات وتتميز الفئات بإمكانية إجراء عمليات الجمع والطرح عليها بمعني أنه يمكن أن تضيف فئة أخري كنوع ومدي الفئة أو نقسم الفئة إلي جزأين ليكون كل قسم منها فئة صغيرة علي سبيل المثال، الفئة العمرية من 16 – 18 سنة يمكن أن تجمع علي فئة العمر 18 – 20 سنة وتصبح فئة واحدة هي 16 – 20 فضلا عن ذلك فإنه يمكن معالجة الفئات معالجات إحصائية متعددة[2].

(1) اعتماد علام، يسرى رسلان، مرجع سابق، ص ص 38 - 39.

(2) حسن محمد حسن، مبادئ الإحصاء الإجتماعى، مرجع سابق، ص 42.

4 – مقاييس الفترة الزمنية والنسبة Interval and Ratio scale

المقياس الفتريscale Interval وهذا المقياس يعد أقوي من السابق حيث هنا يمكن تحديد الفروق بين القيم مثال ذلك درجات الحرارة المئوية (فهرنهيت) ودرجات الاختبار الرقمية: 40،80،65،..... وكذلك عدد ساعات الوقت الإضافي للعمال باعتبارها مقياسا لمستوي التوظف ويؤخذ علي هذا القياس عدم وجود نقطة الصفر المطلق بمعني أن الصفر هنا لا يقيس حالة الانعدام الخاصية وبالتالي لا نستطيع إجراء النسبة بين القيم وأن الطالب الحاصل علي (10) درجات مستواه في التحصيل يساوي خمسة أضعاف أخر حاصل علي (2) درجة[1] وتعتبر بيانات الفترة أكثر أنواع البيانات الإحصائية شيوعا واستخداما في أبحاث العلوم الاجتماعية وهي تعكس القيم الأصلية للظاهرات كأعمار السكان، وكميات الإنتاج الزراعي والصناعي، أعداد السيارات، مساحات المزارع ومساحات البيئات الحضرية درجات الحرارة، وكميات الأمطار[2].

– المقياس النسبى Ratio. ويعد أقوى مستويات القياس بما يمسح بإجراء النسب بين قيم المتغيرات مثال ذلك الأوزان والأطوال ودرجات الحرارة والسرعة[3].

– وعلى خلاف ما ذهبت إليه بعض الكتابات فى الفصل بين مقياس النسبة. من أمثال هنكل Hinkle وآخرين، فإننا نتفق مع ما ذهب إليه بلالوك Blalock من عدم الفصل بين نوعى المقياس حيث يعلل ذلك تعليلا منطقيا حين يرى أنه من الصعوبة بمكان أن نجد مقياسا للفترة لا يكون فى نفس الوقت مقياس نسبة لان الواقع الامبريقى يشير إلى ضرورة وجود الوحدات القياسية أو المعيارية للقياس فلا يعقل أن نجد مادة بلا طول أو كتلة أو نجد درجة حرارة بلا وحدة قياس للحرارة وهى إما درجة مئوية يطلق عليها Centigrade مْ أو درجة فهرنهيت 5F Fahrenheit وتستخدم تلك المقاييس فى حالات تتطلب

(1) مصطفى زايد، مرجع سابق، ص 25.

(2) فتحى عبد العزيز أبو راضى، مرجع سابق، ص 10.

(3) مصطفى زايد، مرجع سابق، ص 25.

قياس الفروق أو المسافات الحقيقية بين قيم معينة وهذه خاصية تجعل مقياس الفترة والنسبة أرقى فى المستوى المقياسى من المقاييس السابقة لكى تؤدى تلك المقاييس وظيفتها. فلو كان المطلوب قياس الفروق والمسافات يستخدم مقياس الفترة (الفئوي)[1].

ويتميز مقياس النسب أو المعدلات Ratio بكل الخصائص التى يتصف بها مقياس الفئات من قدره على وضع البيانات فى ترتيب معين فضلا على ذلك فهو يشتمل على الصفر المطلق، وهذه الخاصية تجعل من الممكن استخدامها فى إجراء كل العمليات الحسابية من جمع وطرح وضرب وقسمة بسهولة تامة. وعلى سبيل المثال، يمكن القول بسهولة ويسر أن الـ 1000 جرام تزيد على 600 جرام بمقدار400جرام وأنها ضعف الـ 500 جرام فهذه الأرقام الصفرية لا تحتاج منها إلى استخدام آلات قاسية حسابية لتحديد العلاقة فيما بينها. كما انه من الممكن استخدام هذا المقياس فى حساب النسبة المئوية الخاصة بكل قيمة من القيم الواقعة عليه والواقع أن مقاييس المعدلات قليلا ما تستخدم فى مجال العلوم الاجتماعية ولكنها تستخدم فى ميدان العلوم الطبيعية فى قياس الأوزان والأطوال والوقت.

ولكى نوضح هذه النقطة نقول أن متغيرات كثيرة تستخدم فى مجال العلوم الاجتماعية مثل النوع والعمر والحالة التعليمية لا تتضمن بالضرورة صفرا فى قياسها بينما متغيرات قياس الأوزان والأطوال تتضمن ذلك الصفر فالكيلو 1000 جرام والمتر 100 سم وهكذا. وفى مجال المعالجات الإحصائية للبحوث الاجتماعية غالبا ما نميل إلى استخدام الفئات الصفرية مثل 10 – 20، 20 – 30 لكى نيسر العمليات الحسابية بدلا من استخدام الفئات غير الصفرية مثل 3 – 6، 6 – 9..... وهكذا[2].

ومن خصائص مقاييس الفترة والنسبة بالإضافة للخصائص التى ذكرناها فى المقياسين السابقين، توحد نوع وحدة القياس فلا يمكن أن نقيس الفرق بين درجتين من الحرارة إحداهما بالفهرنهيت والأخرى بالدرجة المئوية بل يكون الفرق بين درجتين حراريتين مثل

(1) اعتماد علام، يسرى رسلان، مرجع سابق، ص ص 39 – 40.

(2) حسن محمد حسن، مبادئ الإحصاء الإجتماعى، مرجع سابق، ص ص 42 – 43.

38 درجة مئوية، 30 درجة مئوية أى من نفس جنس وحدة القياس. ومن جهة أخرى، إذا قلنا أنه توجد وحدات قياسية لمقياس الفترة، ففى العلوم الاجتماعية قد يتعذر تحقيق ذلك، فمثلا توجد وحدات قياسية أو معيارية لقياس الذكاء، السلطة، الهيئة الاجتماعية والتى نجدها متكررة دائما فى الموضوعات الاجتماعية والتنفسية المختلفة الفترة والخاصية الثانية لمقياس الفترات والنسبة إمكانية استخدام العمليات الحسابية المختلفة من جمع وطرح وضرب وقسمة للدرجات فى عمليات تحليل البيانات فمثلا يمكن إضافة دخل الزوجة إلى الزوج أو إلى دخل باقي أفراد الأسرة. والخاصية الثالثة لمقياس الفترة إذ يهتم بخاصية تساوى الفروق بين المستويات المختلفة مثال ذلك تقسيم الدرجة الواحدة على مقياس الحرارة (الترمومتر) إلى تدريج مقسمة إلى خمسة أقسام يمثل كل جزء منها (2). ومن الدرجة مثلا. ويطلق على هذا النوع من مقاييس الفترة مقياس الفترات المتساوية.Equal intervals Scale

ولكى يتم تدريج فترات متساوية كما قلنا فى مثال مقياس الحرارة يلزم نحدد موضع نقطة مطلقة أو ما نسميه بالاختيار التعسفي لنقطة على المقياس ينسب إليها ترتيب تدرج القيم تصاعديا وبفروق ثابتة على أساس وحدة القياس النوعية المستخدمة. ويطلق علي تلك النقطة نقطة الصفر ومن ثم يطلق علي المقياس فى هذه الحالة مقياس النسبة Ratio Scale حيث يمكن باستخدام النسب تدرج القيم والقول بان القيمة كذا اكبر مرتين أو ثلاث مرات عن القيمة الأخرى المعلومة[1].

ويتبين لنا أنه كلما زاد مستوى القياس للمتغيرات، أى زادت الدقة فى القياس كلما أمكن استخدام مقاييس وأساليب إحصائية على درجة أفضل .والثانية هى أن المتغيرات بمستوى قياس معين يكون التعامل معها بالأساليب الإحصائية المخصصة لهذا المستوى من القياس، كما أنه يمكن أيضا استخدام الأساليب الإحصائية المخصصة لمستويات القياس الأقل[2].

(1) اعتماد علام، يسرى رسلان، مرجع سابق، ص 40،

(2) مصطفى زايد، مرجع سابق، ص 26.

الفصل الثالث

العينـــات

مقدمة.

أولا: تعريف العينة.

ثانياً: أسلوب اختيار العينة (أنواع العينات).

ثالثاً: شروط اختيار العينة.

رابعاً: الاعتبارات التى تدعو إلى استخدام العينات.

خامساً: إطار المعاينة.

سادساً: مصادر الخطأ فى العينات.

سابعاً: العوامل التى تحدد حجم العينة

ثامناً: الأساليب الإحصائية لتحديد حجم العينة.

تاسعاً: التحليل الاحصائى باستخدام العينات.

مقدمه

إن الإجابة على التساؤلات التى يضعها الباحث أو تحقيق الفروض التى يطرحها فى بحثه يتطلب قيامه بجمع بيانات يحصل عليها من ميدان الدراسة، ثم يقوم بعد ذلك بتحليل هذه البيانات واستخلاص النتائج التى قد تؤكد صحة تلك الفروض أو تدحضها والواقع أن البيانات التى يحتاجها الباحث ما هى فى الغالب الأعم إلا ردود وإجابات الناس على أسئلة توجه إليهم ليكشف الباحث بواسطتها عن قيمهم واتجاهاتهم إزاء قضايا ومواقف معينة.

ودراسة المجتمعات الإحصائية تعتمد أساسا على أخذ كل مفردات المجتمع للتعرف على خصائص ومعالم هذا المجتمع وبصفة عامة فإن معالم أي مجتمع (وهي مقادير ثابتة للمجتمع الواحد ولكنها تتغير من مجتمع إلى آخر) هى التى تعطي لهذا المجتمع صفاته دون غيره ونظرا لوجود صعوبات كثيرة تحول دون دراسة جميع مفردات المجتمع بواسطة أسلوب الحصر الشامل، فإننا نجرى دراستنا على جزء صغير من هذا المجتمع أو ما يسمى بالعينة Sample حيث أنه من غير العملي أن يقوم الباحث بالحصول على بيانات من جميع أفراد المجتمع ولكنه يقوم بالحصول على تلك البيانات من قطاع صغير منه وهو ما تعارف عليه علماء الإحصاء بأنه " العينة ".

أولاً: تعريف العينة

هى جزء أو شريحة من المجتمع تتضمن خصائص المجتمع الأصلي الذي نرغب فى التعرف على خصائصه ويجب أن تكون تلك العينة ممثلة لجميع مفردات هذا المجتمع تمثيلا صحيحا[1].

والعينة هى جزء من المجتمع ونقوم بدراستها للتعرف على خصائص المجتمع التى سحبت منه هذه العينة – ولكى تصلح النتائج التى نحصل عليها للتعبير عن المتجمع لا بد وان تكون العينة ممثلة للمجتمع (أي جميع المفردات المراد بحثها) تمثيلا صحيحا[2].

واستخدام العينات معروف منذ القدم ونشاهد له أمثلة عديدة فى الحياة العملية فالكيميائي فى معمله يقوم بدراسة خواص المادة من واقع عينة من هذه المادة والطبيب يقوم بتحليل دم المريض من واقع عينة صغيرة تتكون من بضعه نقاط من دمه ..... الخ[3].

ويتم إتباع دراسة العينات وأسلوب المعاينة وذلك اختصارا للوقت وتوفيرا للجهد والنفقات ولرفع مستوى العمل البحثى وجعله أكثر دقة وذلك لأن دراسة عدد قليل من المفردات أو الحالات يتيح للباحث فرصة جمع معلومات دقيقة وكثيرة عن كل مفردة أو حالة[4].

(1) اعتماد علام، يسرى رسلان، أساسيات الإحصاء الإجتماعى، دار الثقافة للنشر والتوزيع، ص 287.

(2) فاروق عبد العظيم، مختار الهانسى، محمد على محمد، مبادئ الإحصاء، دار المعرفة الجامعية، ص 9.

(3) فاروق عبد العظيم، وآخرون، مرجع سابق، ص 9.

(4) فتحى عبد العزيز أبو راضى، مبادئ الإحصاء الإجتماعى، دار المعرفة الجامعية، ص 17.

ثانياً: أسلوب اختيار العينة

هناك أساليب مختلفة لاختيار العينات ولكن نوع العينة وإجراءات سحبها من المجتمع الإحصائي تختلف من موقف لآخر والاعتبار الجوهري الذى يراعيه الباحث هو الحصول على عينة مناسبة. والواقع أن المعيار الأساسي لكون العينة مناسبة هو أن تحظى العينة برضاء الباحث. بعض الباحثون يلجأون إلى أصدقائهم وجيرانهم وأقاربهم وزملائهم ويعتبرونهم كأفراد ضمن العينة. ويوجد عدة أساليب يعتمد عليها الباحث لاختيار العينات منها[1]:-

(1) العينات اللاحتمالية: Non probability sampling فى تلك الحالات لا تعتمد طريقة اختيار العينة على الأسلوب العشوائى نظرا لأن مجال تطبيقاتها امبريقياً يعتمد على اختيار شريحة أو قطاع معين بطريقة مقصودة. ومن أنواع العينات اللاحتمالية العينة المقصودة والعينة بالحصة.

أ – العينة المقصودة:

إن مجال استخدام هذا النوع من العينات فى الدراسات الاستطلاعية سواء من خلال المقابلات أو الاستبيان بهدف التعرف على اتجاهات فئة معينة من فئات المجتمع حول انتشار وباء معين أو نحو برنامج تليفزيوني أو إذاعي معين وما إلى ذلك وفي هذه الحالة يقتصر الباحث فى اختياره على حي معين من أحياء القاهرة مثلا ثم يقوم الباحث بعد ذلك باختيار عدد من الأسر بهذا الحي دون أي اختيار عشوائي وهنا تبرز أول عيوب العينة اللاحتمالية وتتمثل فى صعوبة تعميم النتائج سواء على مستوى القاهرة كمدينة أو حتى التعميم على مستوى حى معين آخر. أما العيب الثاني فيتمثل فى صعوبة حصول الباحث على تقدير صحيح للخطأ المتوقع بسبب المجازفة[2].

(1) حسن محمد حسن، أساسيات الإحصاء وتطبيقاته، دار المعرفة الجامعية، 1992، ص 29.

(2) اعتماد علام، مرجع سابق، ص 306.

ب - اختيار العينة بالحصة: Quota sampling

وفيها يتم اختيار المبحوثين بنسبة توزيعهم فى المجتمع الاحصائى مثال اختيار20٪ من الإناث 40 ٪ من الذكور وهكذا. ولكن الاختيار الاعتباطي والاختيار بالحصة يعد اختيارا غير اهتمامي، بمعنى أنه لا يوفر فرصة متكافئة لكل مفردات المجتمع الاحصائى لتظهر فى العينة مما يؤدى إلى إخفاق العينة فى أن تمثل المجتمع ككل وتستخدم أحيانا فى المسوح اللاحتمالية للرأي العام وتكون فى هذه الحالة أشبه بالعينة الطبقية. ففي هذه الحالة يعطي القائم بالمقابلة حصة معينة يجب استيفاء بياناتها كأن يلتزم بعدد كبير من الإناث فمن يزيد أعمارهن عن أربعين عاما وأيضا يلزم بعدد كبير من الأشخاص تقل دخولهم السنوية عن (300) جنيه. أو أن يخصص له نسبة معينة من الأطباء فى مجتمع ما وهكذا بحيث يكون الباحث قادرا على أن يتم الحصة المطلوبة منه[1].

(2) العينات الاحتمالية: Probability Samples

لقد طور العلماء أساليب المعاينة الاحتمالية لتجنب المخاطر التى تترتب على اختيار عينة غير ممثلة لمجتمع الدراسة وهذه المخاطر لا يمكن تجنبها تماما ولكن هذه الأساليب تمكننا على الأقل من تحديد نسبة الخطأ المحتمل وتعرف العينة الاحتمالية بأنها العينة التى يتم سحبها بحيث يكون لكل مفردة من مفردات المجتمع فرصة معلومة ومتكافئة فى أن يكون جزءا من العينة.

يتسم هذا النوع من العينات بالخصائص التالية: -

أ - لكل مفردة فى العينة درجة احتمالات معروفة يفترض وجودها بين باقى مفردات تلك العينة.

ب - لجميع مفردات المجتمع الأصلي فرص متساوية للظهور فى العينة.

(1) المرجع السابق، ص 307.

يلزم أن تكون الاحتمالات معروفة لدى الباحث حتى يمكن التوصل إلى الثقل الصحيح للعينة أما إذا لم يعرف الباحث تلك الاحتمالات فإنه قد يستحيل عليه أن يستخدم بنجاح الاستنتاج الإحصائي المعتمد على دلالات بحثية[1].

(3) العينة العشوائية البسيطة: Simple Random sample

العينة العشوائية هي العينة التى تختار بحيث تعطي جميع مفردات المجتمع المراد بحثه نفس الفرصة فى الاختيار وهذا يعنى عدم الاهتمام ببعض المفردات أكثر من البعض الآخر وإتاحة الفرصة المتكافئة أمام كل مفردة للظهور فى العينة ويمكن أن نحقق ذلك بأن نحضر عدا من البطاقات المتشابهة (فى اللون والحجم والوزن وكل شيء) ونكتب على كل بطاقة رقماً يمثل مفرده من مفردات المجتمع وتسحب عددا من هذه البطاقات (بعد خلطها) فنجد أن الأرقام المدونة عليها تعطي لنا المفردات التى تم اختيارها بطريقة عشوائية[2]. وتعرف العينة العشوائية البسيطة بأنها اختياراً بسيطاً بطريقة تتصف بخاصيتين أساسيتين هما: -

أ - أن يتحقق لكل عضو أو مفرده من المجتمع الأصلي درجة احتمال متساوية فى الاختيار.

ب - أن يكون اختيار كل مفردة من مفردات العينة بصورة مستقلة عن الأخرى[3].

لو تصورنا أن أحد الأساتذة بقسم الاجتماع يود إجراء دراسة عن اتجاهات طلاب القسم نحو إدمان المخدرات ثم وضع أسماء هؤلاء الطلاب وعددهم 4000 فى حقيبة كبيرة ثم سحب منها 400 اسم أو أنه أعطى رقما مسلسلاً لكل من هؤلاء الأربعة آلاف طالب تم اختيار 400 رقما من جدول الأرقام العشوائية وقام بعد ذلك باختيار الطلاب الذين يتطابق رقمهم المسلسل مع الأرقام العشوائية المختارة له فإنه يكون بذلك قد أعطى لكل طالب من الطلاب فرصة متكافئة لكى يكون من أحد أفراد العينة.

(1) المرجع السابق، ص 291.

(2) فاروق عبد العظيم وآخرون، مرجع سابق، ص 14.

(3) اعتماد علام وآخرون، مرجع سابق، ص 292.

(4) العينة المنتظمة: Systematic sample:

العينة المنتظمة هي نوع من المعاينة العشوائية بمقتضاها يمكن أن يختار الباحث لو أخذنا فى الاعتبار المثال السابق نسبة ٪10 من عدد الطلاب (400 طالب) ويستطيع الباحث أن يختار هؤلاء الطلاب بطريقة عشوائية فيبدأ بالطالب رقم 8 ثم بعد كل عشر طلاب يقوم باختيار طالب آخر وهكذا أي أنه فى هذه الحالة سيختار الطالب رقم 8، 18، 28، 38 وهكذا. وهذه الطريقة فى الاختيار مقبولة ما لم يكن اختيار الأرقام من البداية يخفض وراءه تحيز الباحث نحو اختيار طلاب بعينهم. والواقع أن الطريقتين السابقتين من طرق اختيار العينات تلائم الباحثين المبتدئين وغيرهم ممن يريدون تجنب التعقيدات الإحصائية وهناك بالإضافة إلى تلك الطرق أساليب أخرى أكثر تطوراً لسحب العينات توفر للعينة صفات أساسية كأن تكون ممثلة ومقبولة ومناسبة من حيث التكاليف[1].

وتعتبر العينة المنتظمة أكثر أفضلية من العينة العشوائية البسيطة وذلك فى حالة توفر قوائم تضم جميع مفردات المجتمع الأصلى غير أن السهولة فى العينة المنتظمة يناظر بعض العيوب من أهمها.

أ – توقع نتائج خاطئة إذا تم استخدام هذا النوع من العينات فى مجتمعات تتسم بتكرار ظواهر دورية.

ب – اقتصار العشوائية فقط في تحديد الرقم الأول فى بداية اختيار العينة[2].

(5) العينات الطبقية: Stratified Samples:

تتميز العينات الطبقية على غيرها من العينات بأنها بالإضافة إلى كونها دراسة للمجتمع ككل فإنها تتيح لنا دراسة كل طبقة من الطبقات على حده وهذا قد يكون مرغوباً فيه فى كثير من الأحيان ففى دراسة لبحث ميزانية الأسرة نحصل على نتائج البحث لكل من الريف

(1) حسن محمد حسن، مرجع السابق، ص 30.

(2) اعتماد علام، مرجع سابق، ص 296.

والحضر على حده وهما الطبقتان اللتان يتكون منهما المجتمع، وبذلك تمكننا العينة الطبقية من دراسة كل من الريف والحضر إلى جانب دراسة المجتمع المصري ككل[1].

تعتمد هذه الطريقة على تقسيم المجتمع الإحصائي إلى فئات أو طبقات ثم اختيار عينة من كل طبقة ففى المثال السابق يمكن لباحث أن يقسم الأربعة آلاف طالب بحسب أصولهم الحضرية إلى طلاب من الدلتا، وطلاب من صعيد مصر، ثم يقوم باختيار عدد من الطلاب الذين ينتمون إلى كل من هذه التقسيمات بطريقة عشوائية ويتحدد عدد الطلاب الذين سيتم اختيارهم من كل طبقة بحسب نسبة تلك الطبقة إلى المجموع الكلي للمجتمع الأصلي فلو فرضنا على سبيل المثال أن ٪50 من جملة عدد الطلاب وهم 4000 طالب، من المدن فإن معنى هذا أن ٪50 من العينة التى حجمها 400 طالب يتم اختيارهم من المدن وهكذا. وعموما يمكن صياغة تلك العلاقة فى القانون التالي:

عدد الأفراد المراد اختيارهم من طبقة معينة =

$$= \text{حجم العينة المراد سحبها} \times \frac{\text{عدد أفراد الطبقة}}{\text{جملة عدد أفراد المجتمع الاحصائى}}$$

فى هذه الحالة من المعتقد أن خطأ المعاينة من المحتمل أن يتناقص ليصل إلى الصفر. فتوزيع الطلاب بحسب موطنهم الأصلي فضلا عما يعكسه من تباين ثقافي بين الطلاب فإنه يقترب كثيراً من الواقع[2].

وتقوم العينة الطبقية على تقسيم المجتمع الأصلي إلى مجموعات يطلق عليها طبقات فرعية أو شرائح Strata ثم نأخذ عينة من كل شريحة على حده بحيث يتكون لدينا عينة ذات حجم كلي (ن) ومن الأهمية بمكان أن يتحدد تعريف الشريحة الطبقية بضرورة ظهور

(1) فاروق عبد العظيم وآخرون، مرجع سابق، ص 17.

(2) حسن محمد حسن، أساسيات الإحصاء وتطبيقاته، مرجع سابق، ص 30.

كل فرد من شريحة واحدة فقط ولا يتكرر فى غيرها. وفى الطريقة البسيطة والشائعة من حيث الاستخدام للعينة الطبقية أن تستخدم فى الاختيار وعند بداية تصميم نموذج العينة الطبقية على الباحث اتخاذ الخطوات التالية:

- حساب تقديري للمتوسطات الحسابية لكل شريحة على حده.
- حساب تقديري للانحراف المعياري لكل شريحة على حده.
- بعد تقدير قيمة (ع) لكل شريحة نبدأ فى وضع أوزان تبعا لحجم الشريحة ونسبة هذا الحجم للمجتمع الأصلي[1].

(6) العينة غير المتناسبة: Disproportionate Sample يلجأ الباحث عادة إلى مثل هذا النوع من العينات إذا كان يريد أن يرفع نسبة عينة جماعة فرعية معينة. فلو أراد الباحث فى مثلنا السابق أن يعرف رأى الطلاب الذين من أصل قروي فى قضية الإدمان لما يتميزون به من وازع ديني وأخلاقي فإنه فى هذه الحالة يزيد من نسبة تمثيل الطلاب القرويين لأن طبيعة مشكلة البحث تقتضي ذلك فيختار الباحث 200 طالب من المناطق الريفية وباقي الطلاب من المدن ومن الصعيد. ولكن فى هذه الحالة ينبغي على الباحث أن يظهر فى تحليله العوامل التى دفعته لمثل هذا النوع من الاختيار.

(7) العينات العنقودية ذات المرحلة الواحدة ومتعددة المراحل Single . stage and Multi. stage cluster Samples

فى حالة العينات كبيرة الحجم يلجأ الباحث إلى هذا الأسلوب من أساليب المعاينة لتخفيض نفقات اختيار العينة والعينة العنقودية ذات المرحلة الواحدة تتمثل فيما يقرره احد الباحثين من اختيار حى سكنى معين من إحدى المدن كعينة للدراسة ثم يختار مجموعة من الأسر التى تقطن ذلك الحي لإجراء مقابلة معهم. معنى هذا أن المقابلات التى سيقوم بها الباحث سوف تتجمع فى حي معين الأمر الذي ساعد على تخفيض الوقت والنفقات

(1) اعتماد علام، يسرى رسلان، مرجع سابق، ص 297.

ونلاحظ هنا أن اختيار العينة تم على مرحلة واحدة.

- أما العينة العنقودية متعددة المراحل فيلجأ إليها الباحث عند اختيار عينة أكبر حجما. فلو أردنا أن ندرس اتجاهات الشباب نحو الإدمان فإنه يمكن أن تحصل على خريطة بأحياء المدينة ثم تختار من بينها عددا من الأحياء الشعبية وعددا أخر من الأحياء الراقية ثم تختار عددا من القطاعات داخل الأحياء وبعد ذلك يتم اختيار من تتم مقابلتهم كأفراد داخل العينة. من ذلك يتضح لنا أن أسلوب العينة العنقودية متعددة المراحل وإن كان يحقق الدقة ويرفع درجة تمثيل العينة للمجتمع الأصلي إلا أنه أسلوب يكتنفه التعقيد ولا يستطيع كثير من الباحثين ذوى الإمكانيات المحدودة الاستعانة به[1].

نظراً لضيق الوقت وكثرة التكاليف والجهود اللازمة لاختيار عينة عشوائية بسيطة فى معظم الأحيان فإننا قد نجرى الاختيار على مراحل متعددة. فإذا كان المجتمع يتكون من أقسام متجانسة نبدأ باختيار بعض هذه الأقسام عشوائيا (كمرحلة أولى) ثم نختار عينة عشوائية بسيطة من كل قسم من الأقسام التي تم اختيارها (كمرحلة ثانية) وقد يحتاج الأمر إلى اختيار عينة عشوائية بسيطة من كل قسم من الأقسام التى تم اختيارها فى المرحلة الثانية و..... وهكذا والعينة التى يتم اختيارها بهذا الشكل تعرف بالعينة متعددة المراحل[2].

---

(1) حسن محمد حسن، أساسيات الإحصاء وتطبيقاته، مرجع سابق، ص ص 29 - 33.

(2) فاروق عبد العظيم وآخرون، مرجع سابق، ص 17.

ثالثاً: شروط اختيار العينة

1 – يجب أن لا تتسم العينة التى تم اختيارها بالتحيز أو المحاباة بمعنى أن تأخذها من بين مفردات المجتمع الأصلي عشوائياً.

2 – أن تكون الظاهرة المراد عمل معاينة لها سائدة ومنتشرة فى المجتمع الأصلي ولا تكون نادرة الحدوث.

3 – يجب أن تكون العينة ممثلة لجميع فئات المجتمع الأصلي.

4 – ضرورة افتراض تجانس مفردات المتجمع الأصلي وفى حالة تعذر ذلك فى بعض المجتمعات غير المتجانسة يلجأ الباحث إلى تقسيمها إلى مجتمعات صغيرة متجانسة.

5 – ضرورة إجراء حصر مسبق لجميع مفردات المجتمع الأصلي المراد بحثه مع تقسيم هذا المجتمع الى وحدات معاينة كل منها داخل قوائم أو ما نسميه إحصائيا بالأطر فعلى سبيل المثال عند دراسة سكان مجتمع ما فإن وحدة المعاينة أما أن تكون الأسرة كوحدة تحليل أو الفرد أو الجماعة وقد يكون المجتمع بالنسبة للمجتمعات الكبيرة.

6 – يجب أن يتناسب اختيار حجم ونوع العينة مع الهدف الأساسي للباحث من العينات مع طبيعة المجتمع أو نوع المشكلة موضوع الدراسة وهكذا[1].

أى أنه يجب أن تتوفر فى العينة الممثلة Representative sample مجموعة من الشروط يمكن تلخيصها فى شرطين أساسين هما:

أ – تكون مفردات العينة ممثلة للمجتمع الذي يجرى عليه البحث تمثيلاً صحيحا وليست ممثلة لمجتمع آخر. بمعني أنه إذا تكررت نفس النتائج على عينات أخرى من نفس

(1) اعتماد علام، يسرى رسلان، مرجع سابق، ص 388.

المجتمع، كانت العينة التى يجرى عليها البحث عينة ممثلة للمجتمع الأصلي أصدق تمثيل، وبذلك يمكن أن تكون خصائص مفردات العينة (إحصائيات العينة) متقاربة أو متشابهة مع خصائص المجتمع (معالم المجتمع) الذي تنتمي إليه.

ب - ألا تكون المفردات المختارة ممثلة لجزء (قطاع) من أجزاء المجتمع الأصلي بل يجب أن تمثل جميع أجزاء المجتمع[1].

(1) فتحى عبد العزيز أبو راضى، مرجع سابق، ص 40.

رابعاً: الاعتبارات التى تدعوا إلى استخدام العينات

يعتبر السبب الرئيسي لاستخدام العينات هو توفير الوقت والجهد والنفقات فإذا كان المال المخصص لإجراء بحث معين أو نوع الباحثين وعددهم أو الوقت اللازم لانجاز هذا البحث لا يسمح بإجراء الحصر الشامل فإننا نضطر لاستخدام العينات لدراسة خصائص المجتمع الذي نجرى البحث لدراسته. وقد تكون هذه العوامل الثلاثة متوفرة لدينا، ومع ذلك نلجأ لاستخدام العينات رغبة فى توفير المال أو اختصاراً للوقت أو ادخاراً للجهد أي بهدف حسن توجيه واستغلال الإمكانيات المادية والفنية. المتاحة فى بعض الأحيان يكون المجتمع الذي ندرسه غير محدد، فإذ أردنا مثلاً فحص إنتاج آلة معينة فالمجتمع هنا يكون ما أنتجته الآلة وما تنتجه الآن وما سوف تنتجه فى المستقبل، لذلك يستحيل فى مثل هذه الحالة إجراء حصر شامل ويكتفى بدراسة عينة من إنتاج الآلة.

قد يؤدي أحيانا فحص المفردات إلى تدميرها فإذا أردنا تحليل الدم لشخص مريض فان الحصر الشامل هنا يعنى سحب كل دم المريض بغرض تحليله، وهذا يعنى قتله، ولذلك لابد فى مثل هذه الحالة من استخدام العينات. أي تجرى التحليل على عينة من بضعة نقاط من دم المريض، وسنجد عموما أنه لابد من استخدام العينات فى الحالات التى يؤدي فيها فحص المفردات إلى إتلافها[1].

اختيار مفردات العينة: -

إن عملية اختيار مفردات العينة من بين مفردات المجتمع الأصلي أو ما يعرف بأسلوب

(1) فاروق عبد العظيم وآخرون، مرجع سابق، ص ص 109-.

سحب العينة من المجتمع كواحدة من المشكلات الخاصة بأسلوب المعاينة، تتوقف أساسا على حجم المجتمع الأصلي. فإذا كان حجم المجتمع صغيرا أي مشتملا على عدد محدد (-fi nite) من المفردات، فإن المشكلة لا تكون مشكلة اختيار العينة من بين مفردات المجتمع، بل تكون مشكلة الحصول على عدد كاف من المفردات لغرض البحث. فمثلا إذا أراد الباحث أن يجرى دراسة على كبار الزراعيين بإحدى القرى، كنموذج لنفس الفئة فى القطر، فقد يحدد هذه الفئة بأنها تشتمل على كل من يمتلك "100 فدانا أو أكثر من الأراضي الزراعية فى القرية " وفى هذه الحالة يكون عدد هؤلاء الملاك قليلا لدرجة أن العينة تستنفذهم جميعا. كما تكون عملية الاختيار من المجتمع الأصلي عملية مشروطة بتحديد المفردات (عدد الملاك) التى تتكون منها العينة المطلوبة وبالطبع كلما كثرت الشروط اللازمة للعينة كلما صعب الحصول عليها وكلما قل عدد المفردات الذين يتم الاختيار من بينهم. أما إذا كان حجم المجتمع الأصلي كبيراً جداً أي مشتملاً على تحدد عدد غير محدد من المفردات المستوفية لجميع الشروط اللازمة فى العينة فإنه من اللازم إجراء عملية اختيار مفردات العينة إما بواسطة الاختيار غير العشوائى (المعاينة العمدية) أو بواسطة الاختيار العشوائى[1].

يستطيع الباحث أن يسلك شتى السبل ويستخدم كافة الأساليب للحصول على عينة للدراسة ولكنه فى كل الأحوال يجب أن يتوخى الحذر من التحيز فى اختيار العينة كما ينبغي عليه أن يتأكد من أن العينة ممثلة لمجتمع الدراسة حتى تكون التعميمات التى يتوصل إليها من تحليلاته مستمرة وقيمة وإلا انعدمت الفائدة من الدراسة[2].

---

(1) فتحى عبد العزيز أبو راضى، مرجع سابق، ص ص 39 - 40.

(2) حسن محمد حسن، أساسيات الإحصاء وتطبيقاته، مرجع سابق، ص 33.

خامساً: إطار المعاينة Sampling Frame

الإطار هو حصر شامل لجميع مفردات المجتمع المراد بحثه فقد يكون الإطار عبارة عن قائمة بالمفردات أو مجموعة من البطاقات أو الخرائط أو.... الخ فعند اختيار العينة يقسم المجتمع إلى أجزاء تسمى وحدات المعاينة (Sampling units) ويكون الإطار عندئذ هو مجموعة القوائم التى تحتوى على هذه الوحدات التى يتكون منها المجتمع. ولما كانت العينات تختار من هذا الإطار وجب أن يكون شاملاً لجميع مفردات المجتمع مع ملاحظة عدم تكرار أي من هذه المفردات لأن عملية التكرار سوف تعطي هذه المفردات فرصة أكبر للاختيار فى العينة وبذلك تتحيز النتائج التى تحصل عليها المفردات التى تكررت فى الإطار ويجب أن يكون الإطار أيضا متجددا حتى تعطى المفردات التى تستجد على الإطار القديم نفس الفرصة فى الظهور فى العينة[1].

ويعتبر إطار المعاينة هو المصدر الذي تؤخذ منه العينة أو بعبارة أخرى هو حصر شامل (القائمة أو الدليل) لجميع مفردات وحدات المجتمع الأصلي المراد دراسته.

مثال ذلك قائمة بأسماء العمال فى أحد المصانع، أو مختلف أنواع الرواسب التى توجد على الشاطئ، أو موقع المحلات العمرانية الريفية على خريطة إحدى الدول. وعند اختيار العينة من المجتمعات المحدودة يقسم المجتمع الأصلي للظاهرة قيد البحث إلى عدة أقسام تسمى وحدات المعاينة (شخص، أسرة، قرية) ويكون إطار المعاينة حينئذ هو عبارة عن القائمة أو مجموعة القوائم التى تتضمن الوحدات التى يتألف منها المجتمع. ويشترط فى إطار المعاينة أن يكون شاملاً لجميع مفردات المجتمع التى يمكن الوصول إليها بسهولة،

(1) فاروق عبد العظيم وآخرون، مرجع سابق، ص، 10

وذلك حتى يكون اختيار العينة سليما. كما يشترط أن يكون إطار المعاينة متجددا حتى تعطي المفردات أو الوحدات التى تستجد على الإطار القديم نفس الفرصة فى الظهور.

ونظرا لأنه فى المجتمعات غير المحددةinfinite يستحيل إجراء حصر شامل لكل مفردات المجتمع فى الوقت المتاح للدراسة، ويكتفي فى هذه الحالة بدراسة عينة بدون تكوين إطار للمعاينة. ويلاحظ على إطار المعاينة وفى مجال الدراسة الجغرافية أنه إما أن يكون إطارا مكانيا Spatial أو غير مكانيا Non – Spatial.

1 – إطار المعاينة المكاني. هو الإطار الذي يكون فيه المكان Location هو الوحدة الرئيسية، كما أنه الأساس فى اختيار العينات التى تمثل التغيرات (الاختلافات) المكانية التى يتميز بها مجتمع الأماكن لمنطقة ما تمثيلاً صحيحاً.

فمثلاً إذا كنا بصدد معاينة خريطة بهدف تحديد مساحة الأراضي التى يشغلها نوعا معينا من النشاط الزراعي على هذه الخريطة، فإننا يجب أن نتأكد من تمثيل كل أجزاء الخريطة تمثيلاً صحيحاً. ويتم ذلك باختيار أحد المعاينات الآتية: –

أ) المعاينة النقطية: Point – sampling أي معاينة نقط تقاطع شبكة مربعات على خريطة المنطقة.

ب) المعاينة الخطية: Line – sampling أي نأخذ عينة من قطاعات عرضية مختلفة على الخريطة.

ج) المعاينة المساحية: Area – sampling أي بأخذ عينة تمثل مساحة مجموعة من المربعات التى تغطي مساحة خريطة المنطقة قيد البحث.

وعلى ذلك يكون إطار المعاينة عبارة عن جميع مفردات المجتمع شكل من أشكال المعاينة الثلاثة.

2 – إطار المعاينة غير المكاني – على الرغم من أن طبيعة عمل الجغرافي عند جمعه

للبيانات ترتبط بإطار المعاينة المكاني، إلا أنه فى بعض الأحيان ولظروف خاصة نجده يهتم بتحديد إطار معاينة غير مكانى ليلائم دراسته فمثلاً إذا كان يصدد اختيار عينة من أسر أحد الأقسام الإدارية فى مدينة ما وذلك لتقدير متوسط الدخل، فإنه يتحتم عليه اختيار عينة من إطار (أو قائمة) تحتوى على جميع أسر هذا القسم الإداري بالمدينة. ولا يجوز له فى هذه الحالة أن يختار العينة من دليل التليفون مثلاً إذ أنه من المعروف أن مثل هذا الدليل لا يتضمن جميع أسر القسم الإداري قيد البحث[1].

(1) فتحى عبد العزيز أبو راضى، مرجع سابق، ص 44.

سادساً: مصادر الخطأ فى العينات

يتضح لنا مما سبق أن خطأ التحيز أمر متوقع لا محالة فى المعاينة الاحتمالية ولا يقتصر هذا التحيز على العينة فقط بل قد نجده أيضا فى عمليات الحصر الشامل حيث تتوافر فرص عديدة للوقوع فى مثل تلك الأخطاء. وقولنا بضرورة وقوع أخطاء يبرره عدم التدريب الكامل للقائمين بالبحث أو المساعدين حول كيفية التغلب على العقبات التى قد تواجههم. هذا فضلاً عن عدم الاستخدام الأمثل للأطر المناسبة والممثلة لاختيار العينة بالطرق الإحصائية السليمة[1].

ويلاحظ أن النتائج التى نحصل عليها من العينة قد لا تماثل تماما النتائج التى نحصل عليها من الحصر الشامل وذلك لأن العينات عرضه لنوعين من الخطأ.

1 - خطأ الصدفة (الخطأ العشوائي) أو ما يسميه البعض بخطأ العينة.

2 - خطأ التحيز.

(1) خطأ الصدفة Random Error

يرجع هذا الخطأ إلى طبيعة الاختيار العشوائي حيث قد تختلف نتائج العينة عن نتائج المجتمع. ويتوقف خطأ الصدفة على كل من حجم العينة وتباين المجتمع وطريقة اختيار العينة وكلما كبرت العينة كلما قل خطأ الصدفة وزادت ثقتنا فى النتيجة، وعلى العكس من ذلك لو زاد تباين مفردات المجتمع لزاد احتمال حدوث الأخطاء العشوائية وعموما لو اختيرت العينة بطريقة عشوائية سليمة لأمكن تقدير هذا النوع من الخطأ من العينة نفسها[2].

(1) اعتماد علام، يسرى رسلان، مرجع سابق، ص 388.

(2) فاروق عبد العظيم وآخرون، مرجع سابق، ص ص 11 - 12.

ويتوقف هذا النوع من الخطأ على درجة تباين المجتمع الأصلي وطريقة اختيار العينة وحجمها فكلما كبر حجم العينة قل خطأ الصدفة وبالتالي زادت درجة الثقة فى النتائج.

هذا ويمكن التحكم فى قيمة هذا الخطأ وتقديره بالطرق الإحصائية وأن كان يصعب تجنب وقوعه إلى حد بعيد. كذلك يجدر الملاحظة أن هذا النوع من الأخطاء يؤثر على العينة وحدها ولا يتأثر به الحصر الشامل بوصفه أحد المصادر الهامة لجمع البيانات.

مثال: فإذا كان لدينا ست أطفال وكانت أعمارهم بالسنة على التوالي 2، 3، 4، 6، 9، 12. أي أن متوسط العمر فى هذه المجموعة

$$= \frac{12 + 9 + 6 + 4 + 3 + 2}{6} = \frac{36}{6} = 6 \text{ سنوات}$$

فإذا سحبنا عينة عشوائية مكونة من حالتين فقط من هؤلاء الأطفال ولتكن 2، 4 فإن متوسط العمر يكون

$$= \frac{6}{2} = 3 \text{ سنوات}$$

وهنا نجد فرقا كبيراً بين متوسط العينة ومتوسط المجتمع الأصلي. وإذا سحبنا عينة أخرى مكونة من حالتين وثالثة، ورابعة لا يكون هذا الاختيار دقيقا إلا فى حالة سحب الحالتين رقم 3، 9 ففى هذه الحالة الأخيرة يمكن القول بأن القيمة المقدرة لأعمار الأطفال تنطبق تماماً على القيمة الحقيقية للأعمار. حيث أن متوسط العينة

$$= \frac{9 + 3}{2} = 6 \text{ سنوات}$$

وهو نفس المتوسط الحقيقي للمجموعة. أى أن خطأ الصدفة يرجع إلى الفرق بين

القيمة المقدرة من العينة والقيمة الحقيقية فى المجتمع الأصلي الذي سحبت منه العينة. ومن هنا لا يستطيع الجزم بأن متوسط القيم فى أية عينة هو نفس المتوسط العام للقيم الحقيقية فى المجتمع الأصلي، فقد يكون عمر أحد أفراد العينة صغيرا فينخفض متوسط العينة وقد يكون كبيراً فيرتفع المتوسط فى العينة عن المتوسط الحقيقي ولا يحدث خطأ الصدفة فى حالة حدوث التعادل. كذلك لا يمكننا الجزم بحدوث هذا التعادل فى أي حالة معينة إذا تركت للصدفة وحدها وكل ما يمكن أن نقوله هنا هو أنه يحتمل حدوث هذا التعادل[1].

(2) خطأ التحيز Bias Error

هذا الخطأ لا يتوقف على عنصر العشوائية أو الصدفة. ويحدث عادة فى اتجاه واحد أي بالزيادة فقط أو بالنقص فقط وتكون خطورته فى أنه لا يمكن حصره أو وضع حدود له.

مثل خطأ الصدفة. وهذا النوع من الخطأ ليس قاصراً فقط على العينات بل قد يتعرض له الحصر الشامل نتيجة لعدم الدقة فى القياس أو عدم كفاءة الباحثين أو غموض كشوف الأسئلة أو إعطاء بيانات غير صحيحة من قبل المبحوثين أو عدم جمع البيانات عن بعض مفردات المجتمع أو جمع البيانات عن بعض مفردات المجتمع أكثر من مرة أو... الخ

وتتعرض العينات لخطأ التحيز لنفس الأسباب التى يتعرض لها الحصر الشامل بالإضافة إلى الأسباب الآتية:

1 - عدم وجود إطار سليم عند سحب العينة، فاستخدم إطار قديم أو إطار غير شامل لجميع مفردات المجتمع يؤدي إلى تحيز العينة للمفردات الموجودة فى الإطار فقط، ولو تكررت بعض المفردات فى الإطار، فإن ذلك يؤدي إلى تحيز العينة للمفردات المتكررة.

2 - حالة عدم إمكانية الوصول لبعض مفردات العينة يستعاض عن هذه الوحدات بوحدات أخرى وذلك قد يؤدي إلى التحيز، ففي حالة عدم تمكن الباحث من الحصول

(1) اعتماد علام، يسرى رسلان، مرجع سابق، ص ص 289 - 290.

على بيانات بعض الأسر نتيجة لتغيبها خارج المسكن نجد أن الاستعاضة قد تؤثر على مدى تمثيل العينة للأسر الصغيرة أو للأسر التى تشتمل على زوجات عاملات.

ج – قد ينشأ التحيز نتيجة لعدم إتباع الطرق السليمة فى حساب التقديرات[1].

ويتسم هذا النوع من الخطأ بالتحيز غالبا نحو جانب واحد إما بالزيادة أو النقصان وتزداد أهمية هذا النوع من الخطأ كلما كبر حجم العينة حيث تقل فرص الخطأ العشوائي.

ويرجع حدوث أخطاء التحيز لعدد من العوامل نذكر من بينها.

– سوء التقدير وعدم توفر الدقة من جانب الباحث وذلك عند قيامه بعمليات الحصر حيث قد تفوته الدقة الكافية فى حساب المتغيرات وكذلك عدم توفيق الباحث فى صياغة الفروض الصحيحة.

– صياغة أسئلة غامضة وغير واضحة للمبحوثين.

– عدم استجابة بعض مفردات العينة لأسئلة المقياس.

– الاختيار المقصود غير العشوائي لمفردات العينة.

– سوء اختيار العينة وقد يحدث نتيجة لسحب العينة من إطار غير كامل.

– عدم دقة القياس[2].

ويتعرض العمل الإحصائي إلى أنواع كثيرة من الأخطاء أثناء تنفيذه ومنها نوعين رئيسيين من أنواع الأخطاء التى يتعرض لها قياس البيانات والتي من شأنها التأثير على النتائج التى نحصل عليها من العينة وهما أخطاء التحيز والأخطاء الاحتمالية.

وأخطاء التحيز هى الأخطاء الناجمة عن تدخل الباحث فى طريقة اختيار العينة فالمعروف مثلاً أن العينة العشوائية تمثل بشكل كبير خصائص المجتمع الذي سحبت منه

(1) فاروق عبد العظيم وآخرون، مرجع سابق، ص 12 – 13.

(2) اعتماد علام، يسرى رسلان، مرجع سابق، ص 290 – 291.

فإذا اختيرت العينة بطريقة شخصية (أي غير عشوائية) فإن ذلك يؤدي إلى زيادة الأخطاء المتوقعة. كذلك تنشأ هذه الأخطاء نتيجة لتحيز الباحث لوجهة نظر خاصة تجاه القرارات المتخذة، ويحدث عادة خطأ التحيز فى اتجاه واحد أما بالزيادة أو بالنقص

ويمكن أن تعزى أخطاء التحيز لعدة عوامل أهمها:

1 - الاختيار المتعمد (غير العشوائي) للعينة.

2 - استبدال أفراد العينة بمفردات أخرى لعدم تمكن الباحث من الوصول لبعض المفردات الأساسية فى العينة.

ج - سوء التقدير وعد توافر الدقة. فقد لا يوفق الباحث فى التفرقة بين ما هو سبب أو نتيجة أو عدم توفر الدقة فى حصر وحساب المتغيرات المحددة لطبيعة الظاهرة ووضع فروض غير سليمة أما الأخطاء الاحتمالية فهي الأخطاء الناجمة عن احتمالات عدم تماثل النتائج التى نحصل عليها مع خصائص المجتمع. فحتى عندما تؤخذ العينة بالأسلوب العشوائي، فإنه تظل هناك احتمالات أخطاء فى مدى تمثيل العينة لخصائص المجتمع الذي أخذت منه. ومنهم أهم هذه الأخطاء ما يطلق عليه إحصائيا خطأ الصدفة أو الخطأ العشوائي[1].

(1) فتحى عبد العزيز أبو راضى، مرجع سابق، ص 10 - 11.

سابعاً: العوامل التى تحدد حجم العينة

عندما يبدأ الباحث فى التفكير فى إجراء دراسته الميدانية يكون من أهم الأسئلة التى ينبغي أن يجيب عنها ذلك السؤال المتعلق بحجم العينة وهل هو مناسب، كبير، أم صغير والإجابة عن ذلك السؤال تتوقف على عدة عوامل هي:

1 - حجم المجتمع الإحصائي الذي ستسحب منه العينة. حيث يشير إلى مجموع الأفراد الذين سيقوم الباحث بسحب العينة من بينهم، وهؤلاء الأفراد يشكلون جزءا من مجتمع أكبر يعرف بالمجتمع الأصلي. فإذا كان الباحث، على سبيل المثال، يريد أن يجرى دراسة على عينة من طلبة كلية الآداب، فإن عدد هؤلاء الطلبة يمثل المجتمع الإحصائي، فى حين أن عدد طلبة جامعة المنصورة بجميع كلياتها يكون بمثابة المجتمع الأصلي. وبطبيعة الحال من المعقول أن نقرر أنه كلما كان حجم المجتمع الإحصائي كبيراً كلما تطلب ذلك أن يكون حجم العينة كبيراً. وبقدر ما يشكل حجم العينة نسبة كبيرة من المجتمع الإحصائي بقدر ما تكون العينة ممثلة لذلك المجتمع فالعينة التى عدد مفرداتها 40 طالبا من فصل مدرسي عدد طلابه 50 طالبا تعد عينة ممثلة تمثيلاً صادقا لذلك الفصل ولكن هذا العدد لا يعتبر عينة ممثلة لمدرسة عدد طلابها 1000 طالب. وبعبارة أخرى، يعتبر كبر حجم العينة ضماناً لأن تكون العينة ممثلة للمجتمع الإحصائي. وليس معنى هذا أن يزيد الباحث من حجم العينة إلي أن تصبح دراسته الميدانية حصراً شاملاً لكل مفردات المجتمع الأصلي الذي يقوم بدراسته ولهذا يلجأ الباحثون إلى استخدام الأساليب الإحصائية لتحديد الحجم المناسب للعينة التى يقومون بدراستها. فزيادة العينة بعد ذلك الحجم لن يضيف إضافة جوهرية إلى درجة الضبط التى ينبغى أن تتميز بها النتائج بقدر ما يضيف من أعباء وتكاليف وما

يستغرق من وقت.

2 - درجة الاختلاف بين مفردات المجتمع الإحصائي. فإذا كانت درجة الاختلاف كبيرة بين أفراد ذلك المجتمع استدعى الأمر زيادة حجم العينة والعكس صحيح. فعندما يكون هناك تماثل تام بين أفراد المجتمع. كأن يكونوا متفقين على قضية عامة، فإن عينة صغيرة جداً منهم تكفي لكى تمثل المجتمع كله. فلو أننا سألنا. 100 فرد هذا السؤال: هل توافق على عودة الشعب الفلسطيني إلى فلسطين؟ لكان ردهم كافيا للتعبير عن اتجاهات ملايين العرب نحو القضية الفلسطينية، بينما لا يكفى هذا العدد كعينة إذا كان السؤال يقصد منه دراسة اتجاهات الأفراد أو نحو السياسية التعليمية.

3 - نسبة الخطأ المسموح به أو المقبول ودرجة الثقة التى يرغب الباحث فى توافرها فى النتائج التى يصل إليها من دراسته للعينة. حيث تعد درجة الضبط المطلوبة فى التنبؤ الذي يبنى على نتائج دراسة هذه العينة ودرجة الثقة فى هذا التنبؤ من العوامل المحددة لحجم العينة. فإذا كان الباحث يسعى إلى التوصل إلى نتائج موثوق بها ويمكن الاعتماد عليها واستخدامها فى التنبؤ، فإن حجم العينة التى سيقوم بدراستها ينبغي أن يكون كبيراً، ولكن كما قلنا سلفاً، كبر حجم العينة يتطلب وقتاً طويلاً وتكلفة ضخمة، لهذا السبب اعتاد الباحثون أن يقبلوا حجم العينة الذي يستطيعون بنسبة ثقة 95٪ أن يعتمدوا على البيانات التى يوفرها لبحثهم وتساعدهم فى استخلاص نتائج يمكن تعميمها على مجتمع الدراسة[1].

وتتفق أراء كثير من الإحصائيين على أن حجم العينة عينة البحث تتوقف على مجموعة من العوامل تنحصر فى: الغرض من البحث، حجم المجتمع الأصلي، مدي تباين الظواهر المختلفة فى قطاعات المجتمع، ودرجة الدقة المطلوبة فى البحث، البيانات المتاحة التى يمكن استخدامها فى تعميم النتائج، والإمكانيات المادية.

ونظرا لعدم وجود اتفاق بين الباحثين على وضع حد معين على أساس علمي. أو

(1) حسن محمد حسن، مبادئ الإحصاء الإجتماعى، دار المعرفة الجامعية، ص 47 - 50.

إحصائي. يحدد الحجم المناسب أو الأمثل للعينة لكى تمثل المجتمع الذي تسحب منه تمثيلاً جيداً، فإن تقدير حجم العينة على مستوى معظم الدراسات والبحوث - تعتبر واحدة من المشكلات الخاصة بأسلوب المعاينة وتطبيق الأساليب الإحصائية، وفي مجال العمل الإحصائي يوجد اتجاهان عند تقدير حجم العينة.

الاتجاه الأول: يعتمد على الخبرة السابقة للباحث فى هذا المجال، حيث أظهرت خلاصة الخبرات والتجارب أن حجم عينة فى حدود ٪10 إلى ٪15 من حجم المجتمع الأصلي يبدو ملائما فى معظم الدراسات والبحوث. ويتميز هذا الاتجاه فى تقدير حجم العينة بسهولته، كما أنه يفيد بعض الباحثين قليلي الخبرة فى مجال العمل الإحصائي.

الاتجاه الثاني: يرتبط أساساً بنظرية الاحتمال Theory of probability مما يتطلب من الباحث الإلمام بقدر وافر من المعلومات الإحصائية والرياضية حتى يستطيع استخدام الأساليب الإحصائية فى تقدير الحجم الأمثل للعينة.

ويعتمد هذا الاتجاه على تحديد العوامل (المتغيرات) التى يتوقف عليها حجم العينة واعتبارها دلائل رئيسية أو مؤشرات أساسية لهذا الغرض وهو أمر يغفله الاتجاه الأول تماما كما يعتمد هذا الاتجاه على توفير بعض المعلومات عن حجم ومعالم المجتمع الأصلي عن طريق العينات التجريبية أو الاسترشادية.

وتتمثل أهم العوامل والمتغيرات الرئيسية المحددة لحجم العينة فى نسبة الخطأ المسموح به (أو درجة الدقة أو الثقة)، ومعامل التشتت (أو الانحراف المعياري) بين مفردات العينة أو المجتمع أن أمكن، والاختلاف النسبي بين المتوسط الحسابي للعينة ومتوسط المجتمع[1].

---

(1) فتحى عبد العزيز أبو راضى، مرجع سابق، ص 19 - 20.

ثامناً: الأساليب الإحصائية لتحديد حجم العينة

يلجأ الباحثون إلى تحديد حجم العينة باستخدام الأساليب الإحصائية تفاديا لتحديده بطريقة تعسفية تثير الانتقادات وتقلل من أهمية العمل العلمي والجهد الذي يبذله الباحث. ويواجه الباحث احتمالين أساسيين عندما يسعى إلى تحديد حجم العينة إحصائيا.

أ - هو ألا يكون على علم بعدد مفردات المجتمع الإحصائي.

ب - هو أن يكون على علم بعدد مفردات المجتمع الإحصائي.

وأخيراً قد تقترح جهة ما على الباحث أن يجرى دراسته على عدد معين من المبحوثين وفى هذه الحالة يميل الباحث إلى تحديد نسبة الخطأ فى هذه العينة ليتأكد من أهمية البيانات التى سيحصل عليها ومدى تمثيل تلك العينة للمجتمع الذي سحبت منه.

تحديد حجم العينة من مجتمع إحصائي غير معلوم

فى كثير من الأحيان لا يجد الباحث بيانات وافية عن عدد أفراد المجتمع الإحصائي الذين سيسحب من بينهم عينة البحث، وذلك لكبر حجم هذا المجتمع، أو لعدم توافر إحصاءات رسمية عن أفراده وفى هذه الحالة يمكن تحديد حجم العينة المطلوب سحبها من مجتمع إحصائي كبير باستخدام المعادلة الآتية: -

$$\text{حجم العينة (ن)} = \frac{Z^2}{\text{خ}_{\text{م}}^2} \times \text{ف}\,(1 - \text{ف})$$

حيث:

Z: القيمة المعيارية عند مستوى ثقة معين وهى فى جميع أحوال الأبحاث تأخذ أحد رقمين هما:

$Z = 1.96$ عند مستوى دلالة 0.05 أو مستوى ثقة 95٪.

$Z = 2.58$ عند مستوى دلالة 0.01 أو مستوى ثقة 95٪.

$خ_م$: الخطأ المعيارى المسموح به وهو أيضاً فى جميع أحوال الأبحاث يأخذ أحد قيمتين هما:

$خ_م$ = 0.05 عند مستوى ثقة 95٪.

$خ_م$ = 0.01 عند مستوى ثقة 95٪.

ف: هى درجة الاختلاف بين مفردات المجتمع الاحصائى وقد اصطلح العلماء على وضعها بقيمة ثابتة أى أن قيم ف = 0.5 دائماً.

تحديد حجم العينة من مجتمع إحصائي معلوم

عند حساب حجم العينة من مجتمع إحصائي معلوم، بمعنى إننا نعرف عدد الأفراد الذين يتكون منهم ذلك المجتمع، فإننا نتبع الخطوات التالية:

- نحسب حجم العينة على أساس أن حجم المجتمع الإحصائي غير معلوم وذلك بالعملية الحسابية السابقة.

- نقوم بعد ذلك بتصحيح حجم العينة، وذلك باستخدام معادلة تصحيح العينة كالاتي: -

معادلة تصحيح حجم العينة:

$$\text{حجم العينة} = \frac{ن_1}{1 + \frac{ن_1 - 1}{ن}}$$

حيث:

ن1: حجم العينة من مجتمع غير معلوم.

ن: حجم المجتمع الاحصائى.

ومن الملاحظ أن حجم العينة من مجتمع إحصائي معلوم العدد أقل من حجم العينة من مجتمع إحصائي غير معلوم العدد، ولذلك فإن استخدام معادلة تصحيح معامل حجم العينة قد أسهم فى ترشيد حجم العينة المناسب للبحث وإن كان الفرق بين حجمي العينتين ليس كبيراً على ما يبدو.

وفى نهاية الأمر يمكن القول بأن اختيار حجم عينة البحث لم يعد يمثل فى الوقت الحالى مشكلة عويصة. فالحاسب الآلي يمكن أن يقدم لنا مقترحات عديدة بهذا الخصوص، كما أن بعض العلماء قد بذلوا جهداً طيباً فى إعداد جداول جاهزة للتغلب على المشكلات المتعلقة بتلك المسألة من ذلك على سبيل المثال جدول حجوم العينات الذي أعده Hush وزميله Backstorm والذى طوره وأضاف إليه Cole [1].

التحليل الإحصائي باستعمال العينات[2]

البيانات الإحصائية هي الأساس للتخطيط الاقتصادي والاجتماعي ولكل البرامج الإنمائية ولمتخذي القرار. وبدخول عصر العولمة ومع الوضع الراهن للدول النامية أصبحت هناك ضرورة مُلحة ومتزايدة للإحصاءات بوجه عام وللبيانات الاقتصادية والاجتماعية بوجه خاص. واستجابة لهذه الحاجة تسعي، كثيراً من دول العالم النامي إلى النهوض بالعمل الإحصائي إلى المستوى اللازم للوفاء باحتياجات المسئولين عن التخطيط للتنمية الاقتصادية والاجتماعية. كما تبذل جُهوداً كبرى في تدريب الكوادر الوطنية القادرة على القيام بإجراء التعدادات والمسوحات وغيرها من نشاطات جمع البيانات وإجراء

(1) حسن محمد حسن، مبادئ الإحصاء الإجتماعى، مرجع سابق، ص 69.

(2) http://www.arab -api.org/course13/c13_1.htm

التحليل بشكل فعّال.

"فالإحصاء (سواء تعداداً أو مسحاً بالعينة) من حيث اللغة هو الإلمام بكل المفردات التي يشملها المجتمع الذي نريد دراسته ومعرفة أو صاف كل مفردة في هذا المجتمع معرفة دقيقة ومحددة بالأعداد. أما علمياً هو عبارة عن تصوير رقمي للواقع في المجتمعات المطلوبة دراستها (المجتمعات البشرية أو غير البشرية)" مثال ذلك تعداد السكان ومسح ميزانية الأسرة فهو تصوير رقمي لأحوال السكان ومستوى معيشتهم على الترتيب.

وننوه بداية بأنه يمكن تقسيم الدراسات والبحوث من حيث المجال أي من حيث درجة الشمول لمفردات المجتمع الأصلي إلى بحوث شاملة وبحوث بطريقة العينات. فالبحث الشامل هو الذي ندرس فيه حاله جميع أفراد المجتمع موضوع البحث بهذه الطريقة إذا كان الغرض منه هو الحصر وذلك مثل تعداد السكان التعداد الزراعي..الخ. وهذا يتطلب تكلفة كبيرة من الوقت والمال والجهد. إن البحث بطريقة العينة فهو الذي نبحث فيه حاله جزء معين (أو نسبة معينة) من أفراد المجتمع الأصلي ثم نقوم بعد ذلك بتعميم نتائج الدراسة على المجتمع كله بتكلفة أقل كثيراً من البحث الشامل.

ومن أمثلة أهم البحوث بالعينة التي تجري على أرض الواقع تلك البحوث التي تستخدم مسوح ميزانية الأسرة وبُحوث القوى العاملة والتي عادة ما تجريها الحكومات أو المؤسسات الدولية أو الإقليمية. كما تشمل مسوحات التجارة والصناعة والمساكن وأبحاث استطلاع الرأي.

مُميزات البحث بالمعاينة وأهميته

واضح أنه من فوائد البحث عن طريق العينة هو اختصار الوقت والجهد اللازمين لإتمام البحث وبالتالي اقتصاد التكاليف. كما يُمكن الحصول بسهولة على الردود الكاملة الدقيقة إذا ما استخدمنا جُزء من المجتمع الكلي. كما أنه يسهل تتبع غير المستجيبين في حالة البحث بالعينة بينما يكون ذلك صعباً في حالة الحصر الشامل. ويُمكن الحصول على بيانات

أكثر من أفراد العينة، وحجمها وتلخيصها وتحليلها على وجه السرعة.

كما تُساعدنا بحوث العينات لمعرفة الدقة التي نتجت عن إجراء حصر شامل والطريقة المثلى هي أن نختار عينة وندرسها دراسة دقيقة وبمقارنة نتائجها مع نتائج التعداد يُمكننا معرفة مدى دقة نتائج الحصر الشامل.

مما سبق يتضح مدى أهمية استخدام العينات والدور الذي تلعبه في الدراسات الكثيرة في مُختلف الميادين، وفي الحقيقة أن استخدام الحصر الشامل أصبح لا يُغني عن استخدام العينة في نفس الوقت فإن تحليل النتائج التي نحصل عليها من تعداد شامل تحتاج إلى وقت طويل وقد تضيع الحكمة من التعداد أو تقل الاستفادة منه إذا ما انتظرنا حتى يتم تحليل النتائج. وفي هذه الحالة يتحتم علينا أن نأخذ عينة ونقوم بتحليل نتائجها لتعطى فكرة عن النتائج النهائية.

أهداف المعاينة

يعد تحديد الهدف الرئيسي للمعاينة أو المشكلة المراد دراستها تحديداً واضحاً، وتحديد أهدافه التفصيلية ربما تكون ذا أهمية كبيرة وذلك لتحديد البيانات المطلوب جمعها واستخدامها من قبل الباحث لكسب ثقة المدى بالبيانات. وبعد ذلك نضع التصميمات المختلفة والممكنة عن طريق الأسئلة المراد الحصول على إجابات عليها. مثلاً يُمكن صياغة أهداف البحث بالسؤال التالي، هل هناك صلة بين التعليم والوعي المصرفي.

إن الغرض الأول من إجراء بحث أو تجربة هو إيجاد إجابات لأسئلة مُعينة حتى نحصل على أساس سليم للتنبؤ عادة ومنه نستطيع اتخاذ إجراء على نتائج العينة فلا بد أن نترجمها ونفسرها بطريقة تُعطي أقصى الفوائد فنوجد التقديرات الإحصائية المختلفة لمعالم المجتمع كما أنه لا بد من قياس دقة هذه التقديرات. وإن من أهم المسائل في تصميم العينات هو الانتهاء إلى معادلة أو معادلات لحساب التقديرات من بيانات العينة وهذه المعادلة أو المعادلات المختارة لا بد أن تحتفظ بكل المعلومات الخاصة بالمجتمع التي حصلنا عليها من

العينة ولا بد من استخدام البيانات لأقصى حد مُمكن.

والتقديرات التي نحصل عليها هي قيم تقريبية لمعالم المجتمع الحقيقية التي نبحث عنها والسؤال المهم هو هل الفرق بين التقدير المحسوب من العينة والقيم الحقيقية للمجتمع صغيراً صغراً كافياً يجعلنا نعتمد على التقدير في دراستنا للمجتمع؟ هنا إذا تم اختيار العينة وحصلنا على التقدير بطرق تعتمد على نظرية الاحتمالات فإنه يُمكننا أن نُقدر دقة هذا التقدير. وإذا كان التقدير يختلف عن القيمة الحقيقية فإن الباحث يُعاني بعض الخسائر إذا ما استخلص نتائجه على أساس هذا التقدير.

وتقديرات معالم المجتمع التي يُمكن الحصول عليها من العينة كثيرة وأبسطها الوسط الحسابي لعينة عشوائية فمن المعروف بأن هذا المتوسط يُعطى تقديراً لمتوسط المجتمع الذي سحبت منه العينة غير أنه لن يكون مُساوياً تماماً لمتوسط المجتمع وذلك يرجع إلى أخطاء المعاينة. ومن التقديرات الأخرى لمعالم المجتمع التي نحصل عليها من المعاينة هي التباين والتفرع والالتواء.

العوامل التى تحدد حجم العينة

عند اختيار عينة من مجتمع الدراسة تثور قضيتان: الأولى تتعلق بحجم العينة والثانية تتصل بالطريقة التى يتم بها سحب العينة وفى هذا الفصل سنهتم فقط بالأساليب الإحصائية لتحديد حجم العينة:

أولاً: العوامل التى تحدد حجم العينة:

1 – حجم المجتمع الاحصائى الذى ستسحب منه العينة.

2 – درجة الاختلاف بين مفردات المجتمع الاحصائى.

3 – نسبة الخطأ المسموح به أو المقبول ودرجة الثقة التى يرغب الباحث فى توافرها فى النتائج التى يصل إليها من دراسته للعينة.

ثانياً: الأساليب الإحصائية لتحديد حجم العينة:

يلجأ الباحثون إلى تحديد حجم العينة باستخدام الأساليب الإحصائية تفادياً لتحديده بطريقة تعسفية تثير الانتقادات وتقلل من أهمية العمل العلمي والجهد الذى يبذله الباحث، ويواجه الباحث احتمالين أساسيين عندما يسعى إلى تحديد حجم العينة إحصائيا:

الأول: هو ألا يكون على علم بعدد مفردات المجتمع الاحصائى.

الثاني: هو أن يكون على علم بعدد مفردات المجتمع الاحصائى.

وأخيراً قد تقترح جهة معينة على الباحث أن يجرى دراسته على عدد معين من المبحوثين وفى هذه الحالة يميل الباحث إلى تحديد نسبة الخطأ فى هذه العينة ليتأكد من أهمية البيانات التى سيحصل عليها ومن مدى تمثيل تلك العينة للمجتمع الذى سحبت منه.

وفيما يلي نتناول أساليب تحديد حجم العينة فى ظل كل احتمال من الاحتمالات السابقة:

1 - تحديد حجم العينة من مجتمع إحصائي غير معلوم

فى كثير من الأحيان لا يجد الباحث بيانات وافية عن عدد أفراد المجتمع الاحصائى الذى سيسحب من بينهم عينة البحث وذلك لكبر حجم هذا المجتمع أو لعدم توافر إحصاءات رسمية عن أفراده وفى هذه الحالة يمكن تحديد حجم العينة المطلوب سحبها من مجتمع احصائى كبير أو غير معلوم باستخدام المعادلة التالية:

$$\text{حجم العينة (ن)} = \frac{Z^2}{\text{خ}_\text{م}^2} \times \text{ف} (1 - \text{ف})$$

حيث:

Z: القيمة المعيارية عند مستوى ثقة معين وهى فى جميع أحوال الأبحاث تأخذ أحد رقمين هما:

$Z = 1.96$ عند مستوى دلالة 0.05 أو مستوى ثقة 95٪.

$Z = 2.58$ عند مستوى دلالة 0.01 أو مستوى ثقة 95٪.

$خ_م$: الخطأ المعيارى المسموح به وهو أيضاً فى جميع أحوال الأبحاث يأخذ أحد قيمتين هما:

$خ_م$ = 0.05 عند مستوى ثقة 95٪.

$خ_م$ = 0.01 عند مستوى ثقة 95٪.

ف: هى درجة الاختلاف بين مفردات المجتمع الاحصائى وقد اصطلح العلماء على وضعها بقيمة ثابتة أى أن قيم ف = 0.5 دائماً.

مثال:

أوجد حجم عينة من مجتمع احصائى غير معلوم إذا علمت أن مستوى الثقة المطلوب توافره فى البيانات هو 95٪؟

الحل:

$$\text{حجم العينة (ن)} = \frac{Z^2}{خ_م^2} \times ف\,(1 - ف)$$

$$\text{حجم العينة (ن)} = \frac{(1.96)^2}{(0.05)^2} \times 0.5\,(1 - 0.5)$$

حجم العينة (ن) = 1536.64 × 0.25 = 384.16 مفردة.

نقرب الكسر لأقرب رقم صحيح فيصبح:

حجم العينة = 385 مفردة.

2 - تحديد حجم العينة من مجتمع إحصائي معلوم

عند حساب حجم العينة من مجتمع احصائى معلوم بمعنى أننا نعرف عدد الأفراد الذين يتكون منهم ذلك المجتمع فإننا نتبع الخطوات التالية:

(أ) نحسب حجم العينة على أساس أن حجم المجتمع الاحصائى غير معلوم من المعادلة التالية:

$$\text{حجم العينة (ن)} = \frac{Z^2}{\text{خ}_{\text{م}}^2} \times \text{ف} \, (1 - \text{ف})$$

حيث:

Z: القيمة المعيارية عند مستوى ثقة معين وهى فى جميع أحوال الأبحاث تأخذ أحد رقمين هما:

Z = 1.96 عند مستوى دلالة 0.05 أو مستوى ثقة 95٪.

Z = 2.58 عند مستوى دلالة 0.01 أو مستوى ثقة 95٪.

خ م: الخطأ المعيارى المسموح به وهو أيضاً فى جميع أحوال الأبحاث يأخذ أحد قيمتين هما:

خ م = 0.05 عند مستوى ثقة 95٪.

خ م = 0.01 عند مستوى ثقة 95٪.

ف: هى درجة الاختلاف بين مفردات المجتمع الاحصائى وقد اصطلح العلماء على وضعها بقيمة ثابتة أى أن قيم ف = 0.5 دائماً.

(ب) نقوم بعد ذلك بتصحيح حجم العينة وذلك باستخدام معادلة تصحيح حجم

العينة كالتالى:

$$\text{حجم العينة} = \frac{ن_1}{1 + \frac{ن_1 - 1}{ن}}$$

حيث:

ن1: حجم العينة من مجتمع غير معلوم كما سيتم حسابها فى الخطوة (أ).

حيث ن: حجم المجتمع الاحصائى.

مثال:

أوجد حجم عينة من مجتمع احصائى حجمه 15000 مفردة إذا علمت أن مستوى الثقة المطلوب توافره فى البيانات هو ٪95؟

الحل:

الخطوة (أ) حساب حجم العينة من مجتمع غير معلوم:

$$\text{حجم العينة (ن)} = \frac{Z^2}{خ_م^2} \times ف\,(1 - ف)$$

$$\text{حجم العينة (ن)} = \frac{(1.96)^2}{(0.05)^2} \times 0.5\,(1 - 0.5)$$

حجم العينة ($ن_1$) = 1536.64 × 0.25 = 384.16 مفردة.

نقرب الكسر لأقري رقم صحيح فيصبح:

حجم العينة ($ن_1$) = 385 مفردة.

الخطوة (ب) تصحيح حجم العينة:

$$\text{حجم العينة} = \frac{ن_1}{1 + \frac{ن_1 - 1}{ن}}$$

$$\text{حجم العينة} = \frac{385}{1 + \frac{385 - 1}{15000}}$$

حجم العينة = 375.24 مفردة

نقرب الكسر لأقري رقم صحيح فيصبح:

حجم العينة = 376 مفردة.

تحديد نسبة الخطأ فى حجم العينة

قد يقرر الباحث إجراء دراسته على عدد معين من الأفراد وفى هذه الحالة التى يحدد فيها الباحث حجم العينة بطريقة تخمينية أو يفرض عليه من الجهة المستفيدة بالدراسة نجده يميل إلى محاولة تحديد نسبة الخطأ فى حجم العينة حتى يطمئن إلى أن البيانات سيحصل عليها والى أن النتائج التى سيتوصل إليها تتمتع بمستوى عالى من الثقة.

وتتحدد نسبة الخطأ فى العينة وفق المعادلة التالية:

$$\text{خطأ العينة} = Z \times \sqrt{\frac{ف\,(1 - ف)}{ن}}$$

حيث:

Z: القيمة المعيارية عند مستوى ثقة معين وهى فى جميع أحوال الأبحاث تأخذ أحد رقمين هما:

$Z = 1.96$ عند مستوى دلالة 0.05 أو مستوى ثقة 95٪.

$Z = 2.58$ عند مستوى دلالة 0.01 أو مستوى ثقة 95٪.

ف: هى درجة الاختلاف بين مفردات المجتمع الاحصائى وقد اصطلح العلماء على وضعها بقيمة ثابتة أى أن قيم ف = 0.5 دائماً.

ن: عدد مفردات العينة.

مثال:

إذا كان لدينا عينة حجمها 600 مفردة سحبت من مجتمع احصائى كبير العدد فما هى نسبة الخطأ المتوقعة فى هذه العينة عند مستوى ثقة بنسبة 95٪ فى البيانات.

الحل:

$$\text{خطأ العينة} = Z \times \sqrt{\frac{\text{ف}\,(1-\text{ف})}{\text{ن}}}$$

$$\text{خطأ العينة} = 1.96 \times \sqrt{\frac{0.5\,(1-0.5)}{600}}$$

خطأ العينة = 1.96 × 0.0204 = 0.04

نسبة الخطأ المعيارى المتوقعة = 0.04 × 100 = 4 ٪

تمارين

1 – أوجد حجم عينة من مجتمع إحصائي حجمه 20000 مفردة إذا علمت أن مستوى الثقة المطلوب توافره فى البيانات هو ٪95؟

2 – أوجد حجم عينة من مجتمع إحصائي حجمه 30000 مفردة إذا علمت أن مستوى الثقة المطلوب توافره فى البيانات هو ٪95؟

3 – أوجد حجم عينة من مجتمع إحصائي حجمه 50000 مفردة إذا علمت أن مستوى الثقة المطلوب توافره فى البيانات هو ٪95؟

4 – إذا كان لدينا عينة حجمها 800 مفردة سحبت من مجتمع احصائى كبير العدد فما هى نسبة الخطأ المتوقعة فى هذه العينة عند مستوى ثقة بنسبة ٪95 فى البيانات.

5 – إذا كان لدينا عينة حجمها 400 مفردة سحبت من مجتمع احصائى كبير العدد فما هى نسبة الخطأ المتوقعة فى هذه العينة عند مستوى ثقة بنسبة ٪95 فى البيانات.

الفصل الرابع

تبويب وعرض البيانات

أولاً: العرض الجدولى للبيانات الإحصائية.

- تبويب البيانات الخام فى جدول تكرارى بسيط.
- تبويب البيانات فى جدول تكرارى ذو فئات.
- تبويب البيانات فى الجدول التكرارى المتجمع الصاعد.
- تبويب البيانات فى الجدول التكرارى المتجمع الهابط.
- الجدول المزدوج.

ثانياً: العرض البيانى للبيانات الإحصائية.

- العرض البيانى للبيانات الغير مبوبة.

1. طريقة الأعمدة البيانية البسيطة.
2. طريقة المنحنى البيانى البسيط.
3. طريقة الخط البيانى المنكسر.
4. طريقة الدائرة البيانية.

5. طريقة الأعمدة البيانية المتلاصقة.

6. طريقة الأعمدة البيانية المجزأة.

- العرض البياني للبيانات الغير مبوبة.

1. المدرج التكراري.

2. المضلع التكراري.

3. المنحنى التكراري.

تبويب البيانات

يقصد بتبويب البيانات عرض هذه البيانات (البيانات الخام) فى جداول مناسبة وذلك حتى يمكن تلخيصها وفهمها واستيعابها واستنتاج النتائج منها ومقارنتها بغيرها من البيانات، كما يسهل الرجوع إليها فى صورة جداول دون الاطلاع على الاستمارات الأصلية التى قد تحمل أسماء أصحابها مما يخل بمبدأ سرية البيانات الإحصائية.

كما يعتبر عرض وتبويب البيانات الإحصائية الخطوة الثانية (بعد تجميع هذه البيانات الخام) في مفهوم التحليل الإحصائي، ويلجأ الباحث إلى حصر وتصنيف هذه البيانات وعرضها بطريقة مختصرة تساعد على فهمها وتحليلها إحصائياً للتعرف عليها ووصفها ومقارنتها بغيرها من الظواهر، والخروج ببعض المدلولات الإحصائية عن مجتمع الدراسة.

عرض البيانات

تتوقف طريقة عرض البيانات على نوع هذه البيانات وعلى الحقائق المطلوب إبرازها. وهناك طريقتان أساسيتان لعرض وتبويب البيانات الإحصائية وهما:

أولاً: العرض الجدولى للبيانات الإحصائية:

بعد عملية تبويب وتعيين الصفات التى تميز المفردات، ترصد النتائج فى جداول مناسبة توضح الشكل النهائى للمجموعات المميزة وتسمى هذه العملية التى يتم تجميع البيانات فى مجموعات مميزة ومتجانسة بعملية التصنيف وتصنف البيانات الإحصائية بوجه عام وفقاً لإحدى القواعد التالية:

1 – تصنيف جغرافى

2 – تصنيف تاريخى أو زمنى.

3 – تصنيف نوعى أو وصفى.

4 – تصنيف كمى.

ويمكن التمييز بين مجموعة أشكال من الجداول الإحصائية نذكرها فيما يلى:

تبويب البيانات الخام فى جدول تكرارى بسيط:

والمقصود بالجدول البسيط هو ذلك الجدول الذى يتم وضع قيم الدرجات فيه مرتبة ترتيباً تصاعدياً فى عموده الأول أما العمود الثانى فيسمى بعمود التكرار ويرصد فيه عدد مرات تكرار كل درجة أو حدث.

مثال:

البيانات التالية هى درجات حصل عليها عشرون طالباً فى مادة الإحصاء الاجتماعي بالفرقة الأولى قسم الاجتماع فى امتحان نهاية العام:

10	12	13	15	10	12	14	15	11	12
15	12	10	12	13	15	12	13	10	14

والمطلوب تبويب هذه البيانات فى جدول توزيع تكرارى بسيط؟

الحل:

يتم ترتيب البيانات دون تكرار تصاعديا ثم وضع هذه البيانات فى العمود الأول من الجدول وتسمى (س) ثم وضع عدد مرات التكرار باستخدام العلامات فى العمود الثانى أما العمود الثالث فيمثل التكرار ويرمز له بالرمز (ك).

س	العلامات	ك
10	////	4
11	/	1
12	~~////~~ /	6
13	///	3
14	//	2
15	////	4
مج		20

مثال:

البيانات التالية هى تقديرات 20 طالباً فى مادة الإحصاء بالفرقة الأولى لقسم الاجتماع فى العام الجامعى 2006/ 2005 والمطلوب هو وضع هذه البيانات فى جدول بسيط؟

جيد	مقبول	جيد	جيد	مقبول	جيد	جيد جداً	مقبول	جيد	جيد جداً
ممتاز	جيد	ممتاز	جيد جداً	مقبول	جيد	ممتاز	جيد	جيد	مقبول

الحل:

التقدير	التكرار
مقبول	5
جيد	9
جيد جداً	3
ممتاز	3
المجموع	20

تبويب البيانات فى جدول تكرارى ذو فئات:

قبل التعرض إلى إعداد هذا الجدول سنقوم أولاً بالتعرف على معنى الفئات وطرق كتابتها.

المقصود بالفئات:

الفئة هى مجموعة من البيانات متشابهة إلى حد كبير جداً فى الصفات، وفى حالة زيادة عدد البيانات الخام التى يتم الحصول عليها من الاستبيان لا يمكن استخدام الجداول البسيطة في التعبير عن هذه الحالات وإلا سنحتاج إلى مئات الصفحات، وإنما يتم تقسيم البيانات إلى مجموعات متقاربة ومتشابهة فى الصفات تسمى فئات.

طرق كتابة الفئات:

يوجد عدة طرق لكتابة الفئات هى:

الطريقة الأولى:

نذكر كلا من الحد الأدنى والحد الأعلى للفئة كما بالجدول التالى:

ف	ك
20 – 10	5
30 – 20	20
40 – 30	50
50 – 40	25

وتنطق الفئة الأولى مثلاً (من 20 إلى 30) وليس (20 شرطة 30) وهذه الطريقة معيبة لأن نهاية الفئة الأولى هى نفسها بداية الفئة الثانية وهكذا وفى هذه الحالة لا نعرف إلى أى فئة ينتمى هذا الرقم.

الطريقة الثانية:

نذكر كلا من الحد الأدنى والحد الأعلى للفئة ولكن نقوم بترك فاصل مقدراه الواحد الصحيح بين نهاية الفئة الأولى وبداية الفئة الثانية وهكذا كما بالجدول التالى.

ف	ك
19 – 10	5
29 – 20	20
39 – 30	50
49 – 40	25

ويعاب على هذه الطريقة أنها لا تصلح فى حالة البيانات التى تحتوى على كسور.

الطريقة الثالثة:

نذكر الحد الأدنى فقط للفئة ونضع بعده شرطة وتنطق الفئة الأولى مثلاً (10 إلى أقل من 20) وهذه الطريقة تصلح لكافة الظواهر.

ف	ك
10 -	5
20 -	20
30 -	50
40 -	25

الطريقة الرابعة:

نذكر الحد الأعلى فقط للفئة ونضع قبله شرطة وتنطق الفئة الأولى مثلاً (أكثر من صفر الى 20) وهذه الطريقة تصلح لكافة الظواهر أيضاً ولكنها أقل شيوعاً.

ف	ك
-20	5
-30	20
-40	50
-50	25

خطوات بناء جدول التوزيع التكراري ذو الفئات:

1 - حساب المدى = أكبر قيمة – أصغر قيمة

2 - حساب عدد الفئات = 3.3 لو (ن)

3 – حساب طول الفئة = المدى / عدد الفئات

4 – اختيار بداية الفئة الأولى أى الحد الأدنى لها مساوى لأقل قيمة موجودة بالبيانات أو أقل بقليل منها فمثلاً تكون من الأرقام الصفرية لتسهيل الحسابات بعد ذلك.

5 – بناء الجدول ووضع العلامات التى تمثل التكرار.

مثال:

قام باحث بجمع بيانات تمثل درجات اختبار مادة الحاسب الآلى لخمسين طالباً من طلاب المرحلة الثانية من الثانوية العامة فى الجدول التالى:

57	42	51	55	70
53	63	47	60	45
55	82	39	65	33
42	65	61	58	64
55	45	53	52	50
39	63	59	36	25
64	54	49	45	65
78	52	41	42	75
26	48	25	35	30
88	46	55	40	20

والمطلوب هو إعداد جدول توزيع تكرارى ذو فئات للجدول السابق؟

الحل:

· المدى = أكبر قيمة – أصغر قيمة = 88 – 20 = 68

• عدد الفئات = 3.3 × لو (ن) = 3.3 × لو (50)

= 3.3 × 1.699 = 5.6

• نقرب عدد الفئات لأقرب رقم صحيح فتكون

عدد الفئات = 7

• طول الفئة = المدى / عدد الفئات = 68 / 7 = 9.7

• نقرب طول الفئة لأقرب رقم صحيح فتصبح

طول الفئة = 10

• نختار بداية الفئة الأولى وهو أصغر رقم = 20

• نبدأ فى بناء الجدول كالتالى:

الفئات	العلامات	التكرار
20 –	////	4
30 –	//// /	6
40 –	//// //// //	12
50 –	//// //// ////	14
60 –	//// ////	9
70 –	///	3
80 – 90	//	2
المجموع		50

تبويب البيانات فى الجدول التكراري المتجمع الصاعد:

ويقصد بالتكرار المتجمع الصاعد هو تجميع تكرار كل فئة على جميع التكرارات السابقة لها بحيث يكون مجموع التكرار التصاعدى للفئة الأخيرة مساوى لمجموع التكرارات.

مثال:

من نفس بيانات المثال السابق كون جدول التكرار المتجمع الصاعد.

الحل:

بنفس الخطوات السابقة نكون جدول التوزيع التكرارى ذو الفئات ومنه نكون جدول التوزيع التكرارى المتجمع الصاعد كالتالى:

حدود الفئات	التكرار المتجمع الصاعد (ك.م.ص)
أقل من 20	صفر
أقل من 30	4
أقل من 40	10
أقل من 50	22
أقل من 60	36
أقل من 70	45
أقل من 80	48
أقل من 90	50

تبويب البيانات فى الجدول التكرارى المتجمع الهابط:

ويقصد بالتكرار المتجمع الهابط هو تجميع تكرار كل فئة على جميع التكرارات التالية لها بحيث يكون مجموع التكرار التنازلى للفئة الأولى مساوى لمجموع التكرارات.

مثال:

من نفس بيانات المثال السابق كون جدول التكرار المتجمع الهابط

الحل:

بنفس الخطوات السابقة نكون جدول التوزيع التكرارى ذو الفئات ومنه نكون جدول التوزيع التكرارى المتجمع الصاعد كالتالى:

حدود الفئات	التكرار المتجمع الهابط (ك.م.هـ)
20 فأكثر	50
30 فأكثر	46
40 فأكثر	40
50 فأكثر	28
60 فأكثر	14
70 فأكثر	5
80 فأكثر	2
90 فأكثر	صفر

الجدول المزدوج

وهو الجدول الذى يربط بين متغيرين فى نفس الوقت وكل متغير منهم له فئاته فيتم بناؤه بإتباع عدة خطوات هى:

1 – تحديد المتغيرين

2 – تحديد المتغير المستقل والمتغير التابع

3 – تحديد فئات كل من المتغيرين

4 - تكوين الجدول بحيث يحتل المتغير المستقل أعلى الجدول أى يكون أفقياً أما المتغير التابع فيحتل الجزء الأسفل أى يكون عمودياً.

5 - وضع العلامات التى تمثل التكرار.

6 - إعادة كتابة الجدول بالأرقام.

مثال:

الجدول التالى يوضح البيانات التى حصل باحث فى دراسة بين النوع و مشاهدة البرامج التعليمية لمجموعة من طلاب الصف الثالث الثانوي على النحو التالى:

النوع	مشاهدة البرامج	النوع	مشاهدة البرامج
ذكر	يشاهد	ذكر	لا يشاهد
ذكر	يشاهد	أنثى	لا يشاهد
أنثى	يشاهد	أنثى	لا يشاهد
ذكر	لا يشاهد	أنثى	يشاهد
أنثى	يشاهد	ذكر	يشاهد
أنثى	لا يشاهد	ذكر	يشاهد
أنثى	لا يشاهد	ذكر	لا يشاهد
ذكر	لا يشاهد	ذكر	لا يشاهد
ذكر	يشاهد	أنثى	يشاهد
أنثى	لا يشاهد	أنثى	لا يشاهد

والمطلوب تكوين الجدول المزدوج للعلاقة بين المتغيرين (النوع ومشاهدة البرامج التعليمية)؟

الحل:

1 – المتغيرين (النوع – مشاهدة البرامج التعليمية)

2 – المتغير المستقل هو النوع والمتغير التابع هو مشاهدة البرامج التعليمية.

3 – فئات المتغير النوع هى (ذكور – إناث)

فئات المتغير مشاهدة البرامج التعليمية (يشاهد – لا يشاهد)

4 – تكوين الجدول بحيث يحتل المتغير المستقل أعلى الجدول أى يكون أفقياً أما المتغير التابع فيحتل الجزء الأسفل اى يكون عمودياً.

كالتالى:

النوع مشاهدة البرامج التعليمية	ذكور	إناث
يشاهد		
لا يشاهد		

5 – وضع العلامات.

النوع مشاهدة البرامج التعليمية	ذكور	إناث
يشاهد	𝍸	////
لا يشاهد	𝍸	𝍸 /

6 – إعادة كتابة الجدول بالأرقام.

النوع / مشاهدة البرامج التعليمية	ذكور	إناث	مج
يشاهد	5	4	9
لا يشاهد	5	6	11
مج	10	10	20

ثانياً: العرض البياني للبيانات الإحصائية

يعتبر العرض البيانى للبيانات الإحصائية بمثابة تلخيص للبيانات الإحصائية فى شكل يسهل منه استيعاب خصائص موضوع بحث الدراسة، وتختلف طرق عرض البيانات المبوبة عن البيانات الغير مبوبة، وسنتعرض لكل منها بالتفصيل فيما يلى: –

أولاً: العرض البيانى للبيانات الغير مبوبة:

والمقصود بالبيانات الغير مبوبة تلك البيانات المفردة أى لا يوجد بها فئات وهناك عدة طرق لعرض البيانات الغير مبوبة.

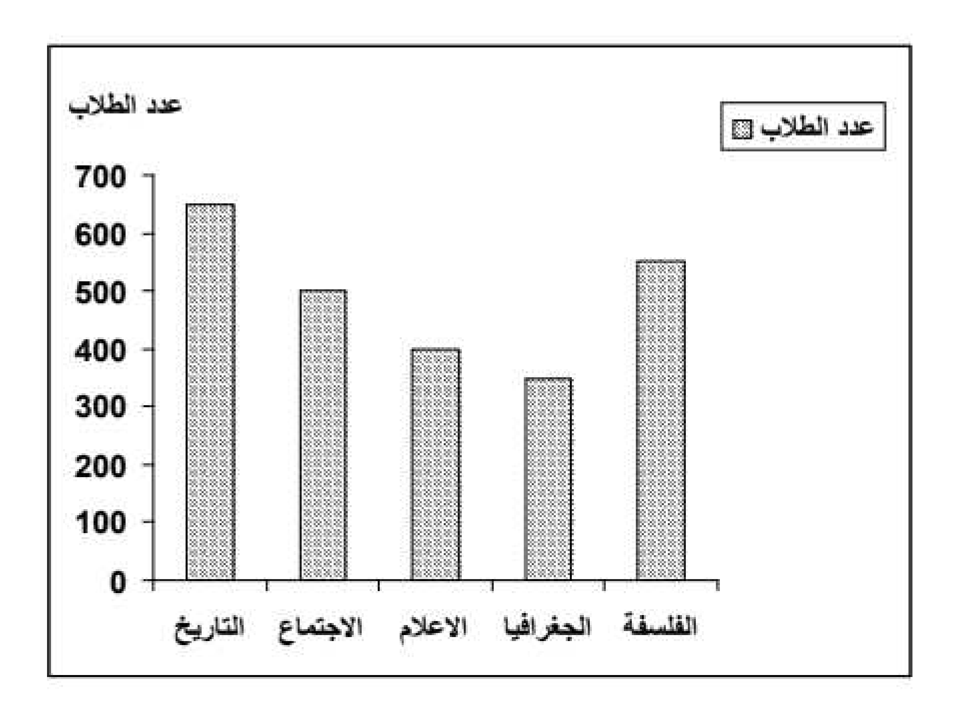
(1) طريقة الأعمدة البيانية البسيطة:

وفى هذه الطريقة يمثل محور السينات قيم المتغير أما محور الصادات يمثل القيمة المقابلة لقيمة المتغير ويتم رسم عمود حول المتغير وارتفاعه يمثل قيمة المتغير.

مثال:

الجدول التالى يوضح أعداد الطلاب ببعض أقسام كلية الآداب جامعة المنصورة والمطلوب عرض هذه البيانات باستخدام طريقة الأعمدة البيانية البسيطة؟

القسم	التاريخ	الاجتماع	الإعلام	الجغرافيا	الفلسفة
عدد الطلاب	650	500	400	350	550



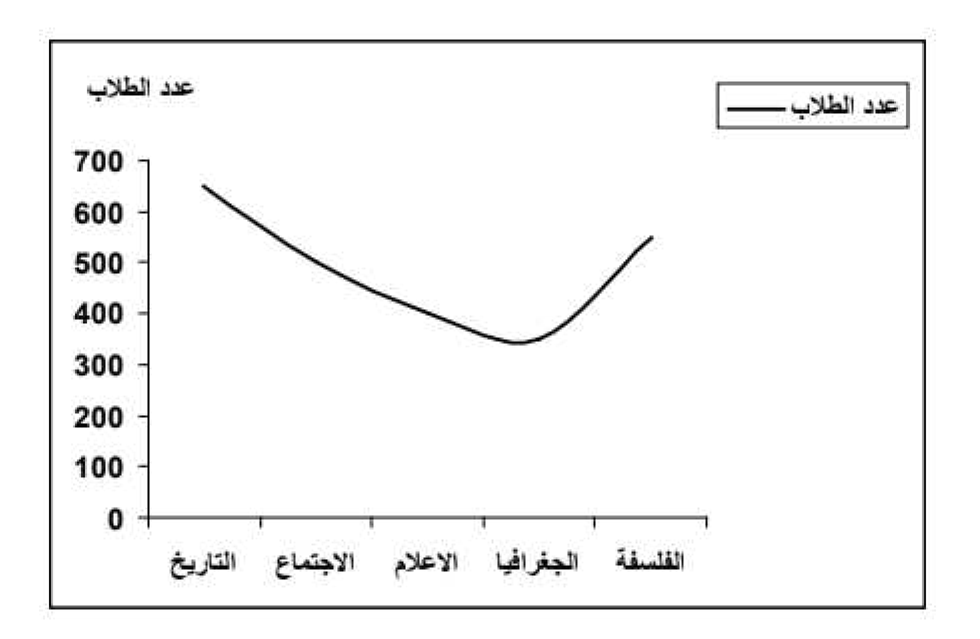
(2) طريقة المنحنى البيانى البسيط:

وفى هذه الطريقة يمثل محور السينات المتغير أما محور الصادات يمثل قيمة المتغير ويتم توقيع نقاط بين كل قيمة من قيم المتغير على محور السينات والقيمة المقابلة على محور الصادات ثم يتم توصيل تلك النقاط بخط منحنى باليد.

مثال:

الجدول التالى يوضح أعداد الطلاب ببعض أقسام كلية الآداب جامعة المنصورة والمطلوب عرض هذه البيانات باستخدام طريقة المنحنى البياني البسيطة؟

القسم	التاريخ	الاجتماع	الإعلام	الجغرافيا	الفلسفة
عدد الطلاب	650	500	400	350	550



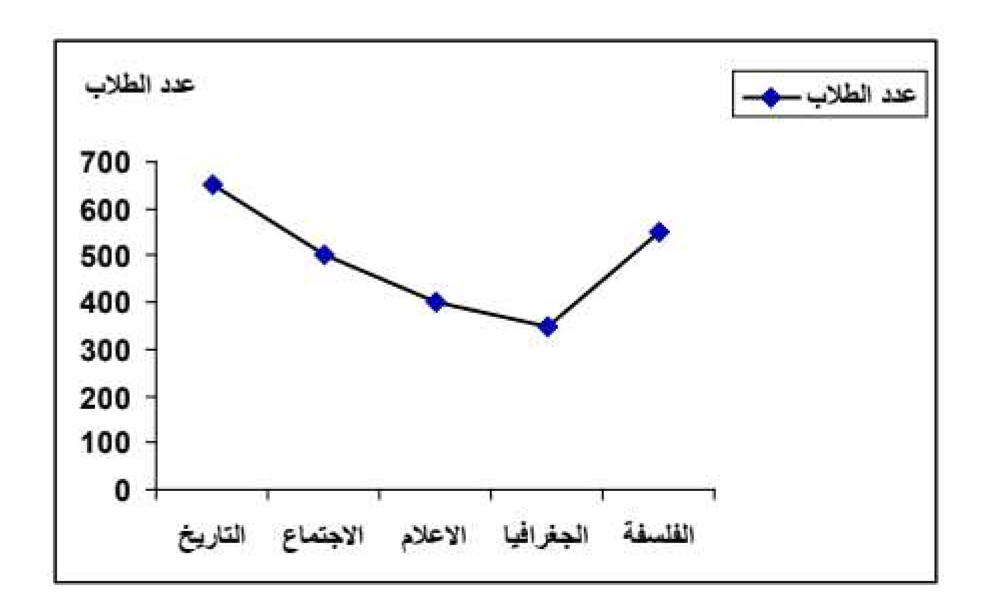
(3) طريقة الخط البيانى المنكسر:

وفى هذه الطريقة يمثل محور السينات المتغير أما محور الصادات يمثل قيمة المتغير ويتم توقيع نقاط بين كل قيمة من قيم المتغير على محور السينات والقيمة المقابلة على محور الصادات ثم يتم توصيل تلك النقاط بخط منكسر باستخدام المسطرة.

مثال:

الجدول التالى يوضح أعداد الطلاب ببعض أقسام كلية الآداب جامعة المنصورة والمطلوب عرض هذه البيانات باستخدام طريقة الخط البياني المنكسر؟

القسم	التاريخ	الاجتماع	الإعلام	الجغرافيا	الفلسفة
عدد الطلاب	650	500	400	350	550



(4) طريقة الدائرة البيانية:

وفى هذه الطريقة يتم رسم دائرة ثم نحسب زاوية قطاع كل قيمة على حدة ونقوم برسم تلك الزاوية داخل الدائرة حتى تنتهى الدائرة.

ونحسب زاوية قطاع الجزء من العلاقة:

$$\text{زاوية قطاع الجزء} = \frac{\text{التكرار الفعلى للجزء}}{\text{مجموع التكرارات}} \times 360$$

مثال:

الجدول التالى يوضح أعداد الطلاب ببعض أقسام كلية الآداب جامعة المنصورة والمطلوب عرض هذه البيانات باستخدام طريقة الدائرة البيانية؟

القسم	التاريخ	الاجتماع	الإعلام	الجغرافيا	الفلسفة
عدد الطلاب	650	500	400	350	550

الحل:

نحسب مجموع التكرارات = 650 + 500 + 400 + 350 + 550

مجموع التكرارات = 2450

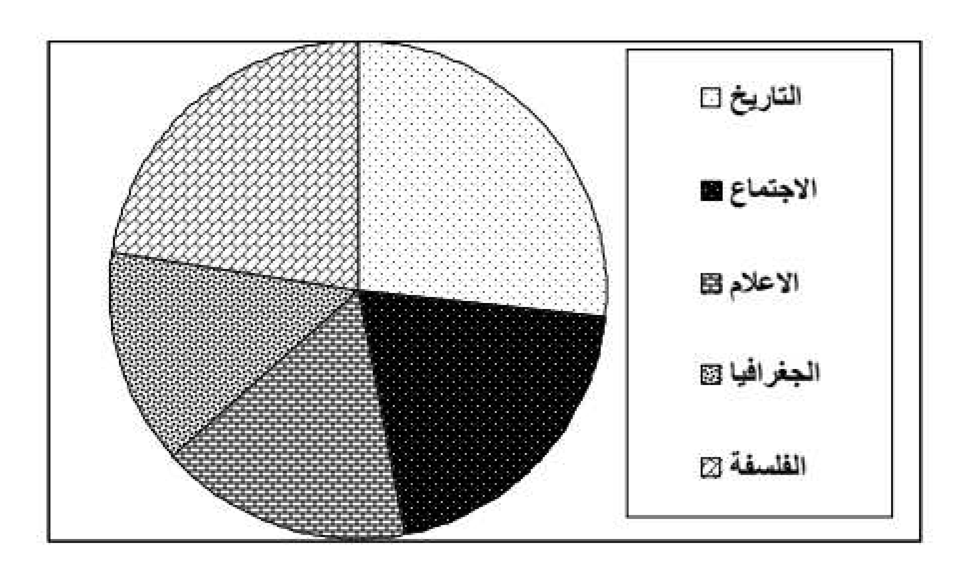
زاوية قطاع التاريخ = $\frac{650}{2450} \times 360 = 95.5^\circ$

زاوية قطاع الاجتماع = $\frac{500}{2450} \times 360 = 73.5^\circ$

زاوية قطاع الإعلام = $\frac{400}{2450} \times 360 = 58.7^\circ$

زاوية قطاع الجغرافيا = $\frac{350}{2450} \times 360 = 51.4^\circ$

زاوية قطاع الفلسفة = $\frac{550}{2450} \times 360 = 80.8^\circ$



(5) طريقة الأعمدة البيانية المتلاصقة:

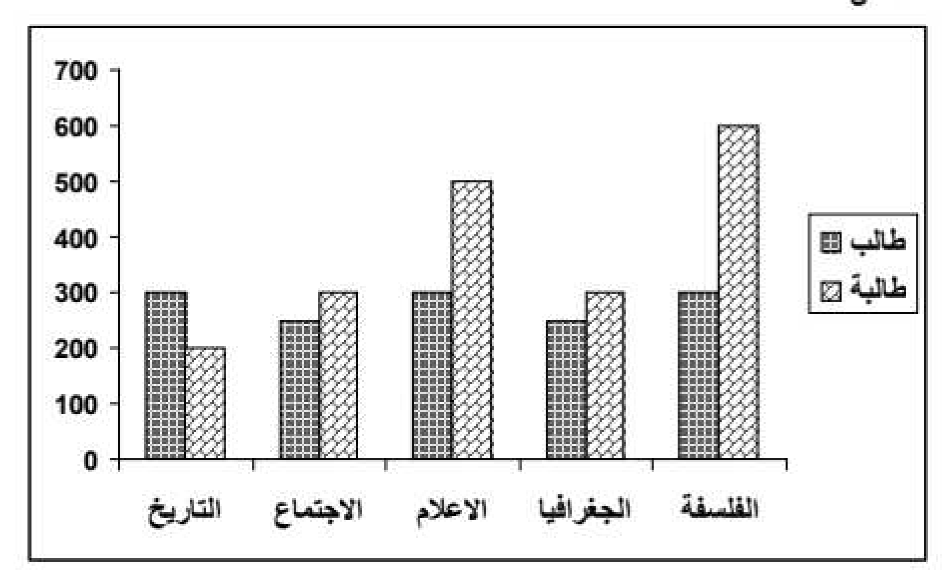
تسمى هذه الطريقة أيضا بطريقة الأعمدة البيانية المتجاورة وهى تشبه طريقة العمدة البيانية البسيطة ولكن يتم رسم عدد من الأعمدة متلاصقة يمثل كل منهم احد قيم المتغير.

مثال:

الجدول التالى يوضح أعداد الطلاب ببعض أقسام كلية الآداب جامعة المنصورة والمطلوب عرض هذه البيانات باستخدام طريقة الأعمدة البيانية المتلاصقة؟

القسم	التاريخ	الاجتماع	الإعلام	الجغرافيا	الفلسفة
طالب	300	250	300	250	300
طالبة	200	300	500	300	600

الحل:



(6) طريقة الأعمدة البيانية المجزأة:

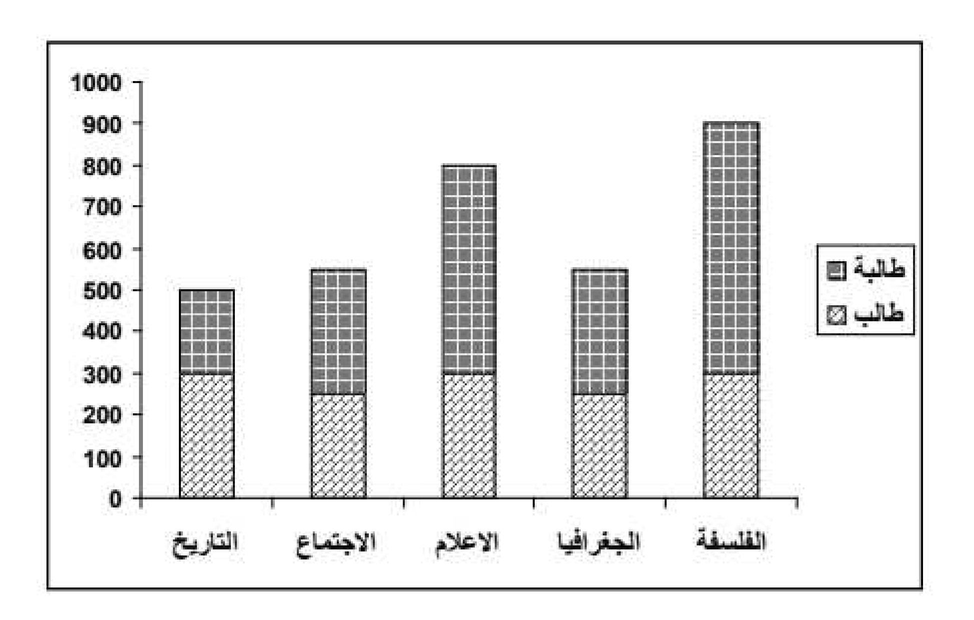
هذه الطريقة تشبه طريقة الأعمدة البيانية البسيطة ولكن يتم رسم عمود يمثل القيمة الأولى للمتغير ثم يليه أو يرتفعه عمود بباقى قيمة المتغير وتكون بادية العمود الثانى هى نهاية العمود الأول.

مثال:

الجدول التالى يوضح أعداد الطلاب ببعض أقسام كلية الآداب جامعة المنصورة والمطلوب عرض هذه البيانات باستخدام طريقة الأعمدة البيانية المجزأة؟

القسم	التاريخ	الاجتماع	الإعلام	الجغرافيا	الفلسفة
طالب	300	250	300	250	300
طالبة	200	300	500	300	600

الحل:



ثانياً: العرض البياني للبيانات المبوبة:

والمقصود بالبيانات المبوبة تلك البيانات المقسمة إلى فئات وهناك عدة طرق لعرض البيانات المبوبة.

(1) المدرج التكرارى:

أحد طرق عرض البيانات المبوبة حيث يتم تخصيص عمود لكل فئة وتكرارها، بحيث يكون طول الفئة هى قاعدة العمود والتكرار هو ارتفاع العمود، ويفضل ترك فراغ كاف قبل الفئة الأولى وفراغ آخر بعد الفئة الأخيرة، أما بالنسبة لمنتصف العمود فيكون هو مركز الفئة.

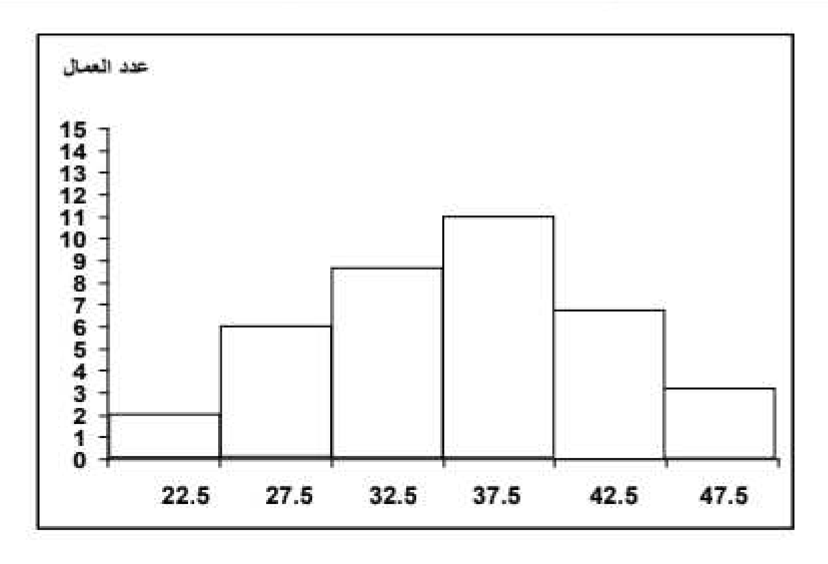
مثال:

اعرض لهذا الجدول بيانياً باستخدام المدرج التكرارى؟

فئات العمر	20 –	25 –	30 –	35 –	40 –	45 –
عدد العمال	2	6	9	11	7	3

الحل:

ف	ك	مركز الفئة
20 –	2	22.5
25 –	6	27.5
30 –	9	32.5
35 –	11	37.5
40 –	7	42.5
45 –	3	47.5



(2) المضلع التكراري:

تخصص لكل فئة وتكرارها نقطة، بحيث يكون الاحداثى السينى لها هو مركز الفئة بينما الاحداثى الصادى لها هو التكرار، نفترض فئة سابقة للفئة الأولى وفئة لاحقة للفئة الأخيرة وتكرار كل منهما صفر، ثم نوصل كل نقطتين متتاليتين بخط مستقيم بالمسطرة.

ملحوظة:

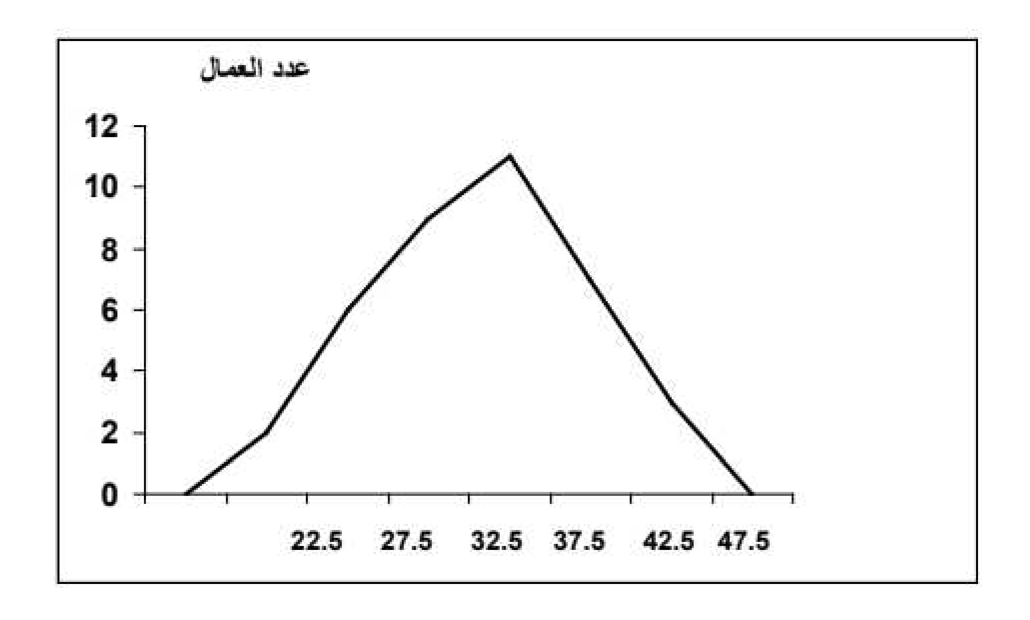
مساحة الشكل تحت المدرج التكرارى = مساحة الشكل تحت المضلع التكرارى.

مثال:

اعرض لهذا الجدول بيانياً باستخدام المضلع التكرارى؟

فئات العمر	20 -	25 -	30 -	35 -	40 -	45 -
عدد العمال	2	6	9	11	7	3

الحل:



(3) المنحنى التكرارى:

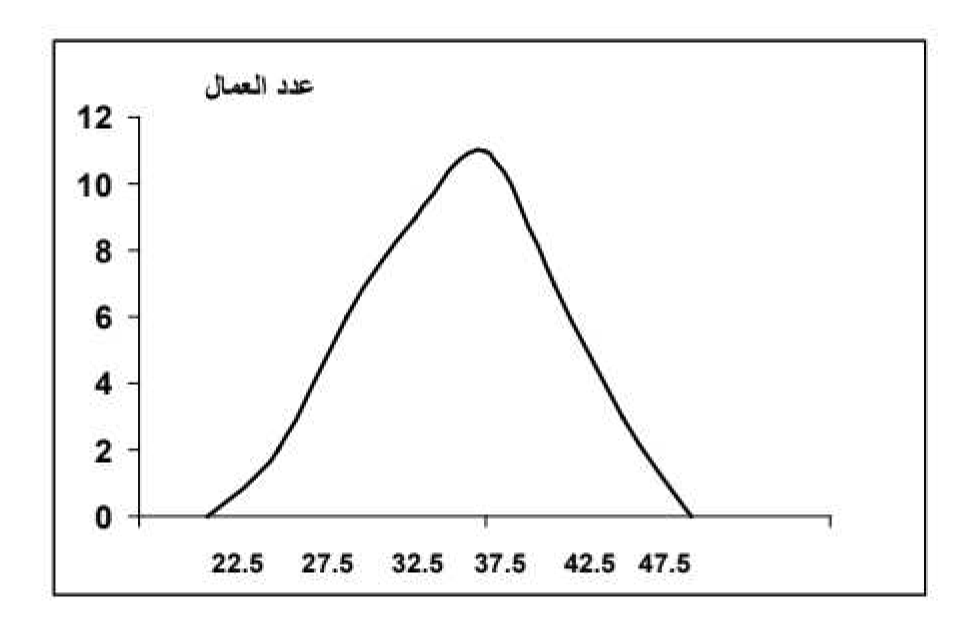
بعد رصد النقاط كما فى الطريقة السابقة نوصل كل نقطتين متتاليتين بمنحنى باليد.

مثال:

اعرض لهذا الجدول بيانياً باستخدام المنحنى التكرارى؟

فئات العمر	20 –	25 –	30 –	35 –	40 –	45 –
عدد العمال	2	6	9	11	7	3

الحل:



تمارين

1 – حصل عدد من الطلاب فى مادة الإحصاء على الدرجات التالية:

5	4	4	5	3	4	2	3	1	2
3	7	4	1	6	3	2	5	3	4
7	3	2	6	5	3	4	2	4	1

المطلوب: تكوين جدول تكرارى بسيط لهذه الدرجات.

2 – تمثل البيانات التالية تقديرات عشرون طالبا فى مادة علم النفس والمطلوب وضعها فى جدول تكرارى بسيط لتلك التقديرات.

ممتاز	مقبول	جيد جدا	مقبول	جيد
جيد جدا	جيد	ضعيف	جيد	مقبول
جيد	ممتاز	مقبول	ضعيف	جيد
جيد جدا	جيد	مقبول	جيد	مقبول

3 – هذه درجات 50 طالبا فى اختبار ذكاء، والمطلوب وضع هذه الدرجات فى جدول تكرارى للفئات.

28	39	33	40	27	55	37	35	37	25
29	28	51	29	51	22	36	44	29	34
32	47	38	25	20	41	36	15	42	33
14	18	34	16	10	46	33	27	27	15
16	27	21	24	17	19	36	19	21	46

4 – الدرجات التالية تمثل درجات 50 طالبا فى أحد الاختبارات:

5	6	5	7	5	6	6	4	5	4
6	6	5	6	6	7	9	8	7	5
5	3	3	5	4	9	7	8	6	7
5	8	8	6	7	7	6	7	7	6
4	6	6	7	6	4	7	7	8	5

والمطلوب: وضع هذه الدرجات فى جدول تكرارى للفئات.

5 – حصل 80 طالبا فى اختبار ذكاء على الدرجات التالية:

46	38	30	20	11	46	23	46	45	18
47	39	33	25	29	49	28	13	36	25
50	43	32	21	19	51	25	15	48	16
49	41	35	27	13	37	29	27	55	37
51	45	21	23	18	50	27	17	12	48
52	42	37	26	14	38	26	14	28	50
53	44	34	22	28	47	30	16	26	36
48	40	31	29	12	35	24	22	20	19

والمطلوب:

- وضع هذه الدرجات فى جدول تكرارى للفئات بحيث يكون عدد الفئات.
- تكوين جدول التكرار المتجمع الصاعد.
- تكوين جدول التكرار المتجمع الهابط.

6 – الجدول التالى يمثل أعداد الكتب بمكتبة الكلية فى مجموعة من التخصصات:

التخصص	علم الاجتماع	علم النفس	التاريخ	اللغة العربية	الجغرافيا
عدد الكتب	550	350	400	600	300

والمطلوب عرض هذه الجدول بيانياً باستخدام الطرق التالية:

- الأعمدة البيانية البسيطة.
- الخط البياني.
- الخط المنكسر.
- الدائرة البيانية.

7 – الجدول التالى يمثل أعداد الذكور والإناث ببعض إدارات أحد الهيئات الحكومية.

الإدارة	الشئون الإدارية	الصيانة	الإحصاء	المعاشات
عدد الذكور	10	20	30	10
عدد الإناث	20	5	60	50

والمطلوب عرض هذه الجدول بيانياً باستخدام الطرق التالية:

- الأعمدة البيانية المتلاصقة.
- الأعمدة البيانية المجزأة.

8 - الجدول التالى يمثل فئات درجات مجموعة من الطلاب فى اختبار للتحصيل وتكراراتهم:

الفئات	5 -	10 -	15 -	20 -	25 -	30 -	35 -	40 -
التكرار	10	13	8	9	12	5	6	7

والمطلوب هو عرض هذا الجدول بيانياً باستخدام الطرق التالية:

- المدرج التكرارى.
- المضلع التكرارى.
- المنحنى التكرارى.

الفصل الخامس

مقاييس النزعة المركزية

أولاً: الوسط الحسابى.

ثانياً: الوسيط.

ثالثاً: المنوال.

رابعاً: العلاقة بين الوسط والوسيط والمنوال.

خامساً: تحديد التواء التوزيع من مقاييس النزعة المركزية.

مقاييس النزعة المركزية

إن الأسلوب البياني في تحليل ودراسة الظواهر لتحديد الخصائص والاتجاهات والعلاقات، يعتمد في دقته على دقة التمثيل البياني نفسه وبذلك ربما تختلف الخصائص من رسم إلى آخر لنفس الظاهرة، وعليه فإنه من الأفضل اللجوء إلى طرق القياس الكمي، حيث يستخدم الباحث الطريقة الرياضية في القياس.

فالهدف الأساسي من استخدام مقاييس النزعة المركزية ومقاييس التشتت هو تلخيص البيانات في محاولة أخرى لوصفها عن طريق التعرف على مركزها ومقدار تشتت البيانات حول هذا المركز (درجة تجانس البيانات) ومن خلال هذين المؤشرين يتمكن الباحث من فهم أبعاد الظاهرة قيد الدراسة.

ومن أهم مقاييس النزعة المركزية التي سنتعرض إليها بالدراسة الوسط الحسابي والوسيط والمنوال، كما سنتعرض بالدراسة لحساب كل منهم من البيانات المفردة (الغير مبوبة) ومن البيانات المبوبة.

أولاً: الوسط الحسابى (المتوسط)

الوسط الحسابي لمجموعة من القيم هو القيمة التى لو أعطيت لكل مفردة فى المجموعة لكان مجموع قيم المفردات الجديدة مساوٍ لمجموع قيم المتغيرات الأصلية.

ويعرف أيضا بأنه مجموع قيم المشاهدات مقسوماً على عددها ويرمز له بالرمز (س/) أو بالرمز (م)

حساب الوسط الحسابى من البيانات الغير مبوبة (المفردة)

يحسب المتوسط الحسابى من البيانات الغير مبوبة من العلاقة التالية:

$$س/ = م = \frac{مجـ س}{ن}$$

حيث: -

س/ = م = الوسط الحسابى

مجـ = مجموع

س = القيمة

ن = عدد الأفراد

مثال: -

احسب الوسط الحسابى لدرجات 8 طلاب في مادة الإحصاء والتي كان بياناتهم كالتالى:

2 – 3 – 5 – 6 – 7 – 8 – 8 – 9

الحل:

$$\text{س/} = \frac{2 + 3 + 5 + 6 + 7 + 8 + 8 + 9}{8} = \frac{48}{8} = 6 \text{ درجات}$$

حساب الوسط الحسابى من البيانات المبوبة

توجد ثلاث طرق لحساب المتوسط الحسابى من البيانات المبوبة هى:

1 - الوسط الحسابى بطريقة مراكز الفئات

$$\text{س/} = \frac{\text{مجـ (س × ك)}}{\text{مجـ ك}}$$

حيث: -

س/ = الوسط الحسابى

مجـ = مجموع

س = مركز الفئة = (بداية الفئة + بداية الفئة التالية) / 2

ك = التكرار

مثال:

الجدول التالى يوضح العلاقة بين فئات الدخل بأحد المصانع وعدد العمال والمطلوب من واقع بيانات الجدول حساب الوسط الحسابى بطريقة مراكز الفئات.

فئات الدخل	100 -	200 -	300 -	400 -	500 -	600 -	700 - 800
عدد العمال	10	12	20	28	16	8	6

الحل:

نكون الجدول التالى:

ف	ك	س	س × ك
100 -	10	150	1500
200 -	12	250	3000
300 -	20	350	7000
400 -	28	450	12600
500 -	16	550	8800
600 -	8	650	5200
700 800-	6	750	4500
مج	100	مج	42600

$$\text{س/} = \frac{42600}{100} = 426 \text{ جنيه}$$

2 - الوسط الحسابى بطريقة الانحرافات

$$\text{س/} = \text{أ} + \frac{\text{مجـ}(\text{ح} \times \text{ك})}{\text{مجـ ك}}$$

حيث: -

س/ = الوسط الحسابى

مجـ = مجموع

ح = الانحراف = س - أ

ك = التكرار

أ = مركز الفئة التى يقابلها أكبر تكرار

مثال:

الجدول التالى يوضح العلاقة بين فئات الدخل بأحد المصانع وعدد العمال والمطلوب من واقع بيانات الجدول حساب الوسط الحسابى بطريقة الانحرافات.

فئات الدخل	100 –	200 –	300 –	400 –	500 –	600 –	700 – 800
عدد العمال	10	12	20	28	16	8	6

الحل:

نكون الجدول التالى:

ف	ك	س	ح	ح×ك
100 –	10	150	300–	3000–
200 –	12	250	200–	2400–
300 –	20	350	100–	2000–
400 –	28	450	صفر	صفر
500 –	16	550	100	1600
600 –	8	650	200	1600
700 – 800	6	750	300	1800
مج	100	مج		2400–

$$\text{س/} = 450 + \frac{-2400}{100} = 450 - 24 = 426$$

3 – الوسط الحسابي بطريقة الانحرافات المختصرة

$$س/ = أ + \frac{مجـ(ح/ \times ك)}{مجـ ك} \times ل$$

حيث: –

س/ = الوسط الحسابى

مجـ = مجموع

ح/ = الانحراف المختصر = (س – أ) / ل

ك = التكرار

أ = مركز الفئة التى يقابلها أكبر تكرار

ل = طول الفئة

مثال:

الجدول التالى يوضح العلاقة بين فئات الدخل بأحد المصانع وعدد العمال والمطلوب من واقع بيانات الجدول حساب الوسط الحسابى بطريقة الانحرافات المختصرة.

فئات الدخل	100 –	200 –	300 –	400 –	500 –	600 –	700 – 800
عدد العمال	10	12	20	28	16	8	6

الحل:

نكون الجدول التالى:

ف	ك	س	ح/	ح/ × ك
100 –	10	150	3–	30–

- 200	12	250	2-	24-
- 300	20	350	1-	20-
- 400	28	450	صفر	صفر
- 500	16	550	1	16
- 600	8	650	2	16
-800 700	6	750	3	18
مج	100	مج		24-

$$\text{س/} = 450 + \frac{24-}{100} \times 100 = 450 - 24 = 426$$

ثانياً: الوسيط

يعرف الوسيط على أنه القيمة التى تتوسط مجموعة من القيم إذا رتبت ترتيباً تصاعدياً أو تنازلياً.

حساب الوسيط من البيانات الغير مبوبة (المفردة)

يعتمد حساب الوسيط من البيانات الغير مبوبة على عدد تلك البيانات فهناك حالتان هما:

(1) إذا كان عدد المفردات فردى (ن فردية)

يوجد رقم واحد يمثل الوسيط ويحسب ترتيبه من العلاقة:

(ن+1) / 2

مثال:

احسب الوسيط من البيانات التالية

20 – 12 – 15 – 10 – 40 – 80 – 61

الحل:

ترتب تصاعدي أولاً:

10	12	15	20	40	61	80

نحسب ترتيب الوسيط = (7 + 1) / 2 = 4، ترتيب الوسيط هو الرابع.

الوسيط = 20.

(2) إذا كان عدد المفردات زوجى (ن زوجيه)

يوجد رقمين يمثلان الوسيط ويحسب عن طريق إيجاد الوسط الحسابى لهما ويحسب ترتيبه من العلاقة:

{ن / 2، ن / 2 + 1}

مثال:

احسب الوسيط من البيانات التالية:

40 – 33 – 20 – 18 – 14 – 15 – 12 – 15

الحل:

نرتب تصاعدي أولاً:

40	33	20	18	15	15	14	12

نحسب ترتيب الوسيط = (8/ 2، 8/ 2 + 1) = (4، 5)، ترتيب الوسيط الرابع والخامس وقيمة الوسيط متوسط القيمتين اللتان ترتيبهما الرابع والخامس.

الوسيط = (15 + 18) / 2 = 16.5.

حساب الوسيط من البيانات المبوبة

يوجد خمس طرق لحساب الوسيط من البيانات المبوبة هى:

1 – الوسيط باستخدام الجدول التكرارى المتجمع الصاعد

$$\text{الوسيط} = \text{الحد الأدنى للفئة الوسيطية} + \frac{\text{ترتيب الوسيط} - \text{ك م ص السابق}}{\text{ك م ص اللاحق} - \text{ك م ص السابق}} \times \text{ل}$$

حيث: -

ترتيب الوسيط = مجـ ك / 2

ك م ص السابق = التكرار المتجمع الصاعد السابق للفئة الوسيطية

ك م ص اللاحق = التكرار المتجمع الصاعد اللاحق للفئة الوسيطية

ل = طول الفئة.

مثال:

الجدول التالي يوضح العلاقة بين فئات الدخل بأحد المصانع وعدد العمال والمطلوب من واقع بيانات الجدول حساب الوسيط باستخدام جدول التكرار المتجمع الصاعد.

فئات الدخل	20 -	30 -	40 -	50 -	60 - 70
عدد العمال	20	40	100	30	10

الحل:

نكون الجدول التالي:

	ف	ك	الحدود الدنيا للفئات	ك م ص	
	20 -	20	أقل من 20	صفر	
	30 -	40	أقل من 30	20	
الحد الأدنى	40 -	100	أقل من 40	60	ك م ص السابق
الحد الأعلى	50 -	30	أقل من 50	160	ك م ص اللاحق
	60 - 70	10	أقل من 60	190	
	مج	200	أقل من 70	200	

ثم نحسب ترتيب الوسيط = 2/ 200 = 100

ثم نبحث داخل عمود (ك م ص) عن القيمتين التى ينحصر بينهما ترتيب الوسيط فنجد أن قيمة ترتيب الوسيط = 100 محصورة بين (60 – 160).

$$\text{الوسيط} = 40 + \frac{100 - 60}{160 - 60} \times 10 = 40 + \frac{400}{100} = 40 + 4 = 44$$

2 – الوسيط باستخدام الجدول التكرارى المتجمع الهابط

$$\text{الوسيط} = \text{الحد الأعلى للفئة الوسيطية} + \frac{\text{ترتيب الوسيط} - \text{ك م هـ اللاحق}}{\text{ك م هـ السابق} - \text{ك م هـ اللاحق}} \times \text{ل}$$

حيث: –

ترتيب الوسيط = مجـ ك / 2

ك م هـ السابق = التكرار المتجمع الهابط السابق للفئة الوسيطية

ك م هـ اللاحق = التكرار المتجمع الهابط اللاحق للفئة الوسيطية

ل = طول الفئة.

مثال:

الجدول التالى يوضح العلاقة بين فئات الدخل بأحد المصانع وعدد العمال والمطلوب من واقع بيانات الجدول حساب الوسيط باستخدام جدول التكرار المتجمع الهابط.

فئات الدخل	20 –	30 –	40 –	50 –	60 – 70
عدد العمال	20	40	100	30	10

الحل:

نكون الجدول التالى:

	ف	ك	الحدود العليا للفئات	ك م ص	
	20 -	20	20 فأكثر	200	
	30 -	40	30 فأكثر	180	
الحد الأدنى	40 -	100	40 فأكثر	140	ك م هـ السابق
الحد الأعلى	50 -	30	50 فأكثر	40	ك م هـ اللاحق
	-70 60	10	60 فأكثر	10	
	مج	200	70 فأكثر	صفر	

ثم نحسب ترتيب الوسيط = 2/ 200 = 100

ثم نبحث داخل عمود (ك م هـ) عن القيمتين التى ينحصر بينهما ترتيب الوسيط فنجد أن 100 محصورة بين (40 – 140)

الوسيط = $50 - \frac{40 - 100}{40 - 140} \times 10 = 50 - \frac{600}{100} = 50 - 6 = 44$

3 - الوسيط بالرسم من الجدول التكراري المتجمع الصاعد

مثال:

الجدول التالى يوضح العلاقة بين فئات الدخل بأحد المصانع وعدد العمال والمطلوب من واقع بيانات الجدول حساب الوسيط بالرسم من جدول التكرار المتجمع الصاعد.

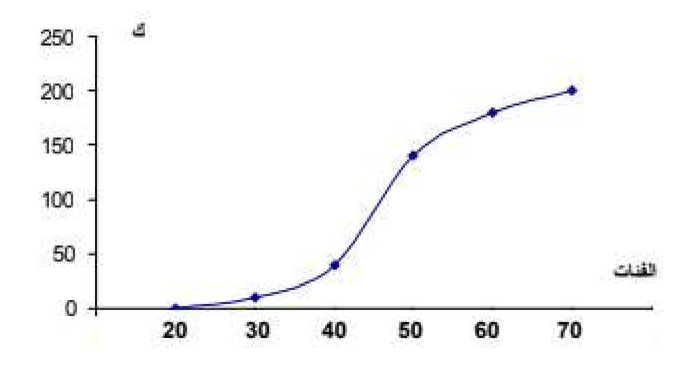
فئات الدخل	20 -	30 -	40 -	50 -	60 - 70
عدد العمال	20	40	100	30	10

الحل:

نكون الجدول التالى:

الحدود الدنيا للفئات	ك م ص
أقل من 20	صفر
أقل من 30	20
أقل من 40	60
أقل من 50	160
أقل من 60	190
أقل من 70	200

ثم نرسم حدود الفئات على محور السينات والتكرار المتجمع الصاعد على محور الصادات ونقوم بتوقيع جميع النقاط ونوصل بينها بخط منحنى باليد كما بالشكل.



ثم نحسب ترتيب الوسيط = مجـ ك / 2 = 2/ 200 = 100 ونوقع هذه النقطة على محور الصادات ونرسم منها خط مستقيم ليقطع المنحنى في نقطة نقوم بإسقاط عمود من نقطة التقاطع ليصل إلى محور السينات لنحصل على قيمة الوسيط عندها.

الوسيط = 44.

4 - الوسيط بالرسم من الجدول التكراري المتجمع الهابط

مثال:

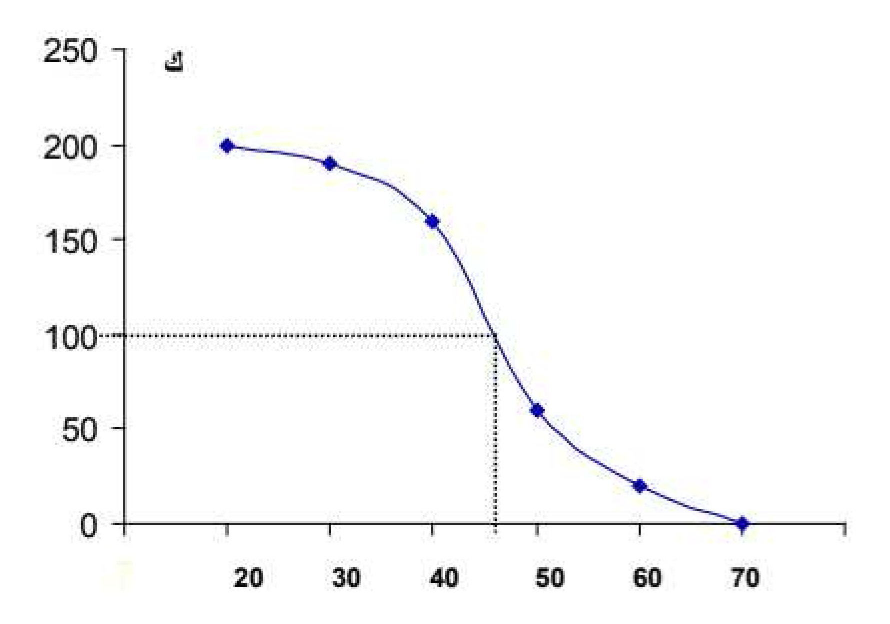
الجدول التالى يوضح العلاقة بين فئات الدخل بأحد المصانع وعدد العمال والمطلوب من واقع بيانات الجدول حساب الوسيط بالرسم من جدول التكرار المتجمع الهابط.

فئات الدخل	20 -	30 -	40 -	50 -	60 - 70
عدد العمال	20	40	100	30	10

الحل:

نكون الجدول التالى:

الحدود العليا للفئات	ك م هـ
20 فأكثر	200
30 فأكثر	180
40 فأكثر	140
50 فأكثر	40
60 فأكثر	10
70 فأكثر	صفر



ثم نحسب ترتيب الوسيط = مجـك / 2 = 2/ 200 = 100 ونوقع هذه النقطة على محور الصادات ونرسم منها خط مستقيم ليقطع المنحنى في نقطة نقوم بإسقاط عمود من نقطة التقاطع ليصل إلى محور السينات لنحصل على قيمة الوسيط عندها.

الوسيط = 44.

5 - الوسيط بالرسم من الجدول التكراري المتجمع الصاعد والهابط معاً

مثال:

الجدول التالى يوضح العلاقة بين فئات الدخل بأحد المصانع وعدد العمال والمطلوب من واقع بيانات الجدول حساب الوسيط بالرسم من جدول التكرار المتجمع الصاعد والهابط معاً.

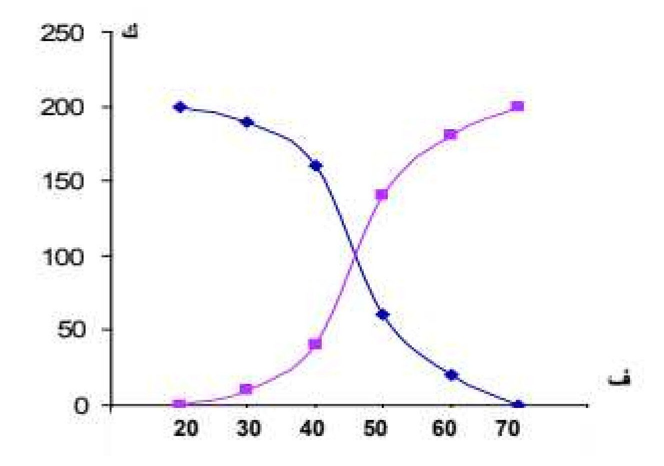
فئات الدخل	20 –	30 –	40 –	50 –	70 – 60
عدد العمال	20	40	100	30	10

الحل:

نكون الجدولين الصاعد والهابط معاً:

الحدود الدنيا للفئات	ك م ص
أقل من 20	صفر
أقل من 30	20
أقل من 40	60
أقل من 50	160
أقل من 60	190
أقل من 70	200

الحدود العليا للفئات	ك م هـ
20 فأكثر	200
30 فأكثر	180
40 فأكثر	140
50 فأكثر	40
60 فأكثر	10
70 فأكثر	صفر



بعد رسم المنحنيين الصاعد والهابط يتقاطعا فى نقطة هذه النقطة لو قمنا بإسقاط عمود منها رأسياً على محور السينات نحصل على قيمة الوسيط = 44.

ولو قمنا برسم خط مستقيم أفقي من نقطة التقاطع ليقطع محور الصادات نحصل على قيمة ترتيب الوسيط = 100.

ثالثاً: المنوال

المنوال هو القيمة الأكثر شيوعاً أو تكراراً.

حساب المنوال من البيانات الغير مبوبة

في حالة تكرار رقم واحد يتم اختياره كمنوال أما فى حالة تكرار رقمين بنفس عدد مرات التكرار يتم اختيارهما معاً كمنوال أما إذا زاد أحدهما عن الآخر يتم اختيار ذو التكرار الأكبر وفى حالة عدم تكرار أي رقم يكون المنوال قيمته لاشيء أو لا يوجد منوال.

مثال: احسب المنوال في كل من الحالات التالية: -

7 – 8 – 9 – 8 – 10 – 8 – 12 المنوال = 8

10 – 12 – 10 – 15 – 12 – 10 المنوال = 10

15 – 16 – 15 – 20 – 16 – 30 المنوال = 15، 16

20 – 30 – 40 – 140 – 50 60– المنوال = لا يوجد

حساب المنوال من البيانات المبوبة

يوجد أربعة طرق لحساب المنوال من البيانات المبوبة طريقتان جبريتان وطريقتان بيانيتان وسنتناولهما بالشرح فيما يلي.

أولاً – المنوال بطريقة الفروق لبيرسون.

$$\text{المنوال} = أ + \frac{ف_1}{ف_1 + ف_2} \times ل$$

حيث:

أ = الحد ألدنة للفئة المنوالية والمقصود بدايتها.

$ف_1$ = ك – ك1

$ف_2$ = ك – ك2

ك = تكرار الفئة المنوالية

ك1 = تكرار الفئة التى تسبق الفئة المنوالية

ك2 = تكرار الفئة التى تلى الفئة المنوالية

ل = طول الفئة

مثال:

أوجد المنوال بطريقة بيرسون من الجدول التالى:

فئات الدخل	10 –	20 –	30 –	40 –	50 –	60 –	70 – 80
عدد العمال	5	12	22	38	22	12	5

الحل:

	ف	ك	
	10 –	5	
	20 –	12	
	30 –	22	ك1
أ	40 –	38	ك
	50 –	22	ك2
	60 –	12	
	70 80–	5	

ثم نحدد الفئة المنوالية من خلال أكبر رقم فى عمود التكرار ثم نحدد الحد الأدنى لهذه الفئة وهو بدايتها وهو أ = 40، ثم نحــدد (ك، ك1، ك2).

نحسب ف1 = ك – ك1 = 38 – 22 = 16

نحسب ف2 = ك – ك2 = 38 – 22 = 16

نحسب ل = 10

ثم نعوض فى القانون:

$$\text{المنوال} = 40 + \frac{16}{16 + 16} \times 10$$

المنوال = 40 + 5 = 45

ثانياً – المنوال بيانيا باستخدام طريقة الفروق لبيرسون.

مثال:

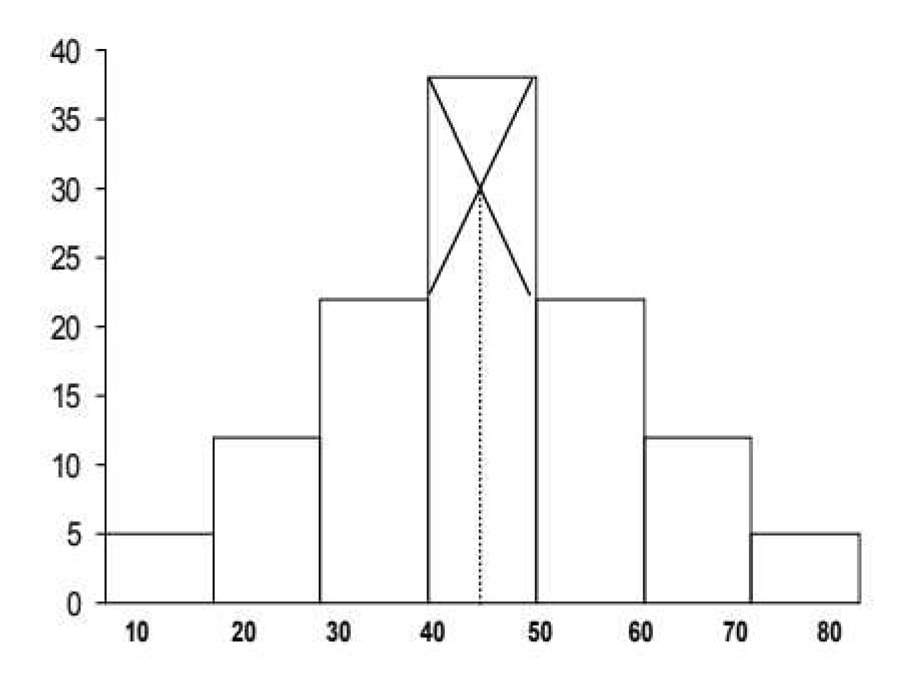
أوجد المنوال بيانياً باستخدام طريقة الفروق لبيرسون من الجدول التالى:

فئات الدخل	١٠ –	٢٠ –	٣٠ –	٤٠ –	٥٠ –	٦٠ –	٧٠ – ٨٠
عدد العمال	٥	١٢	٢٢	٣٨	٢٢	١٢	٥

الحل:

نرسم الجدول السابق بالشكل التالى ثم نبحث عن أطول عمود ونوصل حافتيه بحافتي العمود السابق والتالي فنحصل على تقاطع هو المنوال.

المنوال = 45



ثالثاً: المنوال باستخدام طريقة الرافعة كينج.

$$\text{المنوال} = \text{أ} + \frac{\text{ك}1}{\text{ك}1 + \text{ك}2} \times \text{ل}$$

حيث:

أ = الحد ألدنة للفئة المنوالية والمقصود بدايتها.

ك1 = تكرار الفئة التى تسبق الفئة المنوالية

ك2 = تكرار الفئة التى تلى الفئة المنوالية

ل = طول الفئة

مثال:

أوجد المنوال بطريقة الرافعة كينج من الجدول التالى:

فئات الدخل	10 -	20 -	30 -	40 -	50 -	60 -	70 - 80
عدد العمال	5	12	22	38	22	12	5

الحل:

	ف	ك	
	10 -	5	
	20 -	12	
	30 -	22	ك1
أ	40 -	38	
	50 -	22	ك2
	60 -	12	
	70 - 80	5	

ثم نحدد الفئة المنوالية من خلال أكبر رقم في عمود التكرار ثم نحدد الحد الأدنى لهذه الفئة وهو بدايتها وهو أ = 40، ثم نحـــدد (ك1، ك2).

ك1 = 22

ك2 = 22

نحسب ل = 10

ثم نعوض في القانون:

$$\text{المنوال} = 40 + \frac{22}{22 + 22} \times 10$$

المنوال = 40 + 5 = 45

رابعاً – المنوال بيانيا باستخدام طريقة الرافعة كينج.

مثال:

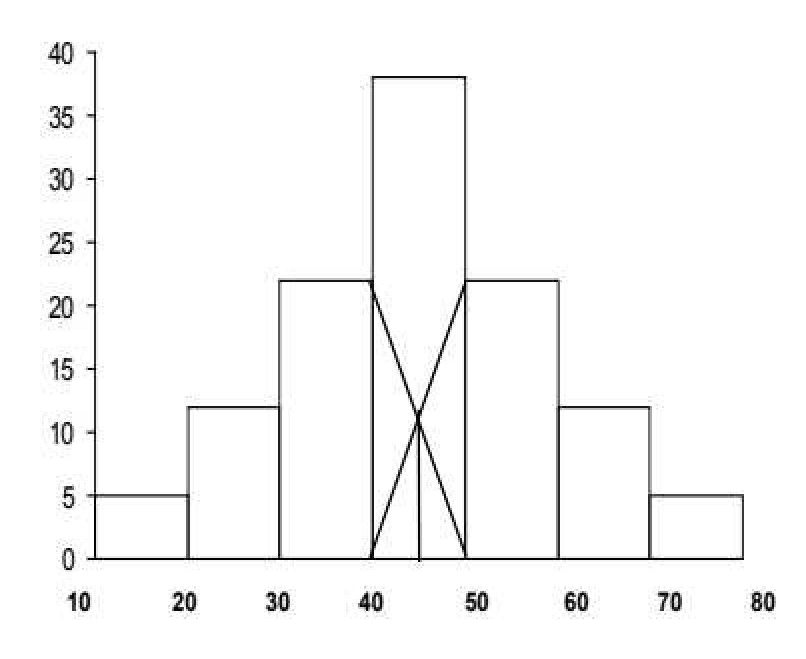
أوجد المنوال بيانياً باستخدام طريقة الرافعة كينج من الجدول التالى:

فئات الدخل	10 –	20 –	30 –	40 –	50 –	60 –	70 – 80
عدد العمال	5	12	22	38	22	12	5

الحل:

نرسم الجدول السابق بالشكل التالى ثم نبحث عن أطول عمود ونصل حافتيه بحافتي العمود السابق والتالي فنحصل على تقاطع هو المنوال.

المنوال = 45



العلاقة بين الوسط والوسيط والمنوال

المنوال = 3 × الوسيط – 2 × الوسط

مثال:

إذا علمت أن قيمة الوسط = 5 وقيمة الوسيط = 10 احسب قيمة المنوال.

الحل:

المنوال = 3 × الوسيط – 2 × الوسط

المنوال = 3 × 10 – 2 × 5

المنوال = 30 – 10 = 20

تحديد التواء التوزيع مباشرة من مقاييس النزعة المركزية

1 – المنحنى معتدل التوزيع:

عندما يكون:

الوسط = الوسيط = المنوال

2 – المنحنى ملتوى التواء موجب:

عندما يكون:

الوسط > الوسيط > المنوال

3 – المنحنى ملتوى التواء سالب:

عندما يكون:

الوسط < الوسيط < المنوال

مثال

إذا علمت أن قيمة الوسط = 5 وقيمة الوسيط = 10 احسب قيمة المنوال، ثم حدد نوع التواء التوزيع.

الحل:

المنوال = 3 × الوسيط – 2 × الوسط

المنوال = 3 × 10 – 2 × 5

المنوال = 30 – 10 = 20

نلاحظ أن

الوسط < الوسيط < المنوال

التوزيع ملتوي التواء سالب.

تمارين

1 – احسب الوسط الحسابى والوسيط للدرجات الخام التالية:

5 ـ 3 ـ 2 ـ 8 ـ 17 ـ 4 ـ 10

من قيمة الوسط والوسيط احسب قيمة المنوال ثم حدد التواء التوزيع.

2 – أوجد الوسط الحسابى والوسيط فى كل حالة من الحالات التالية ومنها أوجد قيمة المنوال ثم حدد التواء التوزيع.

- 7 – 12 – 9 – 11 – 8
- 105 – 107 – 104 – 103 – 102 – 111
- 22 – 23 24 – 20 – 9 – 18 – 35 – 3 – 39 – 36

3 – احسب الوسيط والمنوال لكل حالة من الحالات التالية:

- 5 ـ 8 ـ 2 ـ 4 ـ 9 ـ 10 ـ 6
- 6 ـ 9 ـ 7 ـ 10 ـ 4 ـ 5 ـ 8
- 10 – 12 – 10 – 15 – 12 – 15 – 15 – 20
- 20 – 25 – 30 – 20 – 40 – 60 – 70
- 13 – 15 – 18 – 12 – 10 – 18 – 15 – 18

4 – الجدول التالى يمثل فئات الأجر الأسبوعى لعمال مصنع.

الأجر الأسبوعى	2 –	4 –	6 –	8 –	10 – 12
عدد العمال	10	40	70	50	30

والمطلوب:

- احسب الوسط الحسابي بطريقة مراكز الفئات
- احسب الوسط الحسابي بطريقة الانحرافات
- احسب الوسط الحسابي بطريقة الانحرافات المختصرة
- احسب الوسيط باستخدام جدول التكرار المتجمع الصاعد
- احسب الوسيط باستخدام جدول التكرار المتجمع الهابط
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الهابط
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد والهابط
- المنوال بطريقة بيرسون
- المنوال بيانياً بطريقة بيرسون
- المنوال بطريقة الرافعة كينج
- المنوال بيانياً بطريقة الرافعة كينج

5 – من واقع بيانات الجدول التالي: –

ف	20 –	30 –	40 –	50 –	60 – 70
ك	20	40	100	30	10

• احسب الوسط الحسابي بطريقة مراكز الفئات
• احسب الوسط الحسابي بطريقة الانحرافات
• احسب الوسط الحسابي بطريقة الانحرافات المختصرة
• احسب الوسيط باستخدام جدول التكرار المتجمع الصاعد

· احسب الوسيط باستخدام جدول التكرار المتجمع الهابط

· احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد

· احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الهابط

· المنوال بطريقة بيرسون

· المنوال بيانياً بطريقة بيرسون

· المنوال بطريقة كينج.

6 – من واقع بيانات الجدول التالي: –

ف	100 –	200 –	300 –	400 –	500 –	600 –	700 –
ك	10	12	20	28	16	8	6

· احسب الوسط الحسابي بطريقة مراكز الفئات

· احسب الوسط الحسابي بطريقة الانحرافات

· احسب الوسط الحسابي بطريقة الانحرافات المختصرة

· احسب الوسيط باستخدام جدول التكرار المتجمع الصاعد

· احسب الوسيط باستخدام جدول التكرار المتجمع الهابط

· احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد

· احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الهابط

· احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد والهابط معاً

· المنوال بطريقة بيرسون

· المنوال بيانياً بطريقة بيرسون

· المنوال باستخدام طريقة الرافعة كينج

· المنوال بيانياً باستخدام طريقة الرافعة كينج.

7 – من واقع بيانات الجدول التالي: –

ف	ك
10 –	5
20 –	12
30 –	22
40 –	38
50 –	22
60 –	12
70 – 80	5
المجموع	116

· احسب الوسط الحسابي بطريقة مراكز الفئات

· احسب الوسط الحسابي بطريقة الانحرافات

· احسب الوسط الحسابي بطريقة الانحرافات المختصرة

· احسب الوسيط باستخدام جدول التكرار المتجمع الصاعد

· احسب الوسيط باستخدام جدول التكرار المتجمع الهابط

· احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد

· احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الهابط

· احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد والهابط معاً

- المنوال بطريقة بيرسون
- المنوال بيانياً بطريقة بيرسون

8 – من واقع بيانات الجدول التالي: –

ف	ك
– 5	2
– 10	4
– 15	6
– 20	8
– 25	10
– 30	16
– 35	40
– 40	24
– 45	14
– 50	11
60 – 55	5

- احسب الوسط الحسابي بطريقة مراكز الفئات
- احسب الوسط الحسابي بطريقة الانحرافات
- احسب الوسط الحسابي بطريقة الانحرافات المختصرة
- احسب الوسيط باستخدام جدول التكرار المتجمع الصاعد

• احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد

• المنوال بطريقة بيرسون

9 – من واقع بيانات الجدول التالي: –

ف	ك
40 –	11
50 –	20
60 –	16
70 –	28
80 –	13
90 – 100	12
المجموع	100

• احسب الوسط الحسابي بطريقة مراكز الفئات

• احسب الوسط الحسابي بطريقة الانحرافات

• احسب الوسط الحسابي بطريقة الانحرافات المختصرة

• احسب الوسيط باستخدام جدول التكرار المتجمع الصاعد

• احسب الوسيط باستخدام جدول التكرار المتجمع الهابط

• احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد

• احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الهابط

• المنوال بطريقة بيرسون

• المنوال بيانياً بطريقة بيرسون

10 - من واقع بيانات الجدول التالي: -

ف	ك
100 -	10
200 -	25
300 -	13
400 -	28
500 -	15
700-600	9
المجموع	100

- احسب الوسط الحسابي بطريقة مراكز الفئات
- احسب الوسط الحسابي بطريقة الانحرافات
- احسب الوسط الحسابي بطريقة الانحرافات المختصرة
- احسب الوسيط باستخدام جدول التكرار المتجمع الصاعد
- احسب الوسيط باستخدام جدول التكرار المتجمع الهابط
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الهابط
- المنوال بطريقة بيرسون
- المنوال بيانياً

الفصل السادس

مقاييس التشتت

أولاً: المدى.

ثانياً: التباين والانحراف المعيارى.

ثالثاً: الانحراف المتوسط.

رابعاً: الالتواء وتحديد اعتدالية التوزيع.

مقاييس التشتت

لا تعتبر مقاييس التمركز كافية لوصف مجموعة من البيانات وصفاً كاملاً فقد تتساوى بعض العينات في الوسط الحسابي بالرغم من اختلاف توزيع بياناتها حول مركزها (درجة تجانس البيانات). فالعينات التالية ذات وسط حسابي واحد (8) ولكنها بلا شك تختلف عن بعضها.

عينة 1	8	8	8	8	8
عينة 2	4	3	6	16	11

فالوسط الحسابي يمثل مركز البيانات لكنه لا يبين مدى التفاف أو بعثرة البيانات حول هذا الوسط، ولهذا لا بد من وجود مقياس آخر مع المقاييس المركزية لقياس درجة التجانس أو التشتت في داخل هذه البيانات.

إن الدرجة التي تتجه بها البيانات الرقمية للانتشار حول قيمة وسطى تسمى تشتت أو توزيع البيانات.

ومن أهم مقاييس التشتت المدى والتباين والانحراف المعيارى والانحراف المتوسط.

أولاً: المدى

المدى هو الفرق بين أكبر قيمة وأصغر قيمة.

حساب المدى من البيانات الغير مبوبة

المدى = أكبر قيمة – أصغر قيمة

مثال: احسب المدى للبيانات التالية:

95 – 200 – 250 – 300 – 110 – 90 – 150 – 100 – 350 – 80

الحل:

نرتب القيم أولاً: (80 – 90 – 95 – 100 – 110 – 150 – 200 – 250 – 300 – 350)

المدى = 350 – 80 = 270

حساب المدى من البيانات المبوبة

المدى = الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى

مثال: احسب المدى للجدول التالي:

الفئات	16 –	20 –	24 –	28 –	32 – 36
عدد المبحوثين	10	15	40	20	15

الحل:

المدى = الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى

المدى = 36 – 16 = 20

ثانياً:
التباين والانحراف المعيارى

يرمز للتباين بالرمز ع2

بينما يرمز للانحراف المعيارى بالرمز ع

أي أنه إذا تم حساب أحدهما فيمكن حساب الآخر لأن الانحراف المعيارى هو جذر التباين.

التباين من البيانات الغير مبوبة

هناك طريقتان لحساب التباين من البيانات الغير مبوبة:

الأولى: باستخدام القانون العام من الدرجات الخام كالتالي

$$ع^2 = \frac{مجـ س^2}{ن} - \left(\frac{مجـ س}{ن}\right)^2$$

مثال:

احسب التباين والانحراف المعيارى للقيم التالية ومنه احسب الانحراف المعيارى لكل من المتغيرين س، ص على حده.

س	23	21	19	19	18
ص	19	19	18	14	15

الحل:

نكون الجدول التالى:

س	$س^2$	ص	$ص^2$
23	529	19	361
21	441	19	361
19	361	18	324
19	361	14	196
18	314	15	225
100	2016	85	1467

ثم نعوض فى القانون العام لحساب التباين:

بالنسبة للمتغير (س)

$$ع^2 س = \frac{مجـ س^2}{ن} - \left(\frac{مجـ س}{ن}\right)^2$$

$$ع^2 س = \frac{2016}{5} - \left(\frac{100}{5}\right)^2 = 3.2$$

وبالتالي فان قيمة تباين المتغير س = $ع^2$= **3.2**

ومنها فان قيمة الانحراف المعيارى = جذر التباين

$$ع = \sqrt{3.2} = 1.78$$

بالنسبة للمتغير (ص)

$$ع^2 ص = \frac{مجـ ص^2}{ن} - \left(\frac{مجـ ص}{ن}\right)^2$$

$$ع^2 ص = \frac{1467}{5} - \left(\frac{85}{5}\right)^2 = 4.4$$

وبالتالي فان قيمة تباين المتغير ص = ع2= 4.4

ومنها فان قيمة الانحراف المعيارى = جذر التباين

$$ع = \sqrt{4.4} = 2.1$$

الثانية: باستخدام الطريقة المختصرة «طريقة الانحرافات»

$$ع^2 = \frac{مجـ ح^2}{ن}$$

حيث ح هو الانحراف = س – م س

مثال:

احسب الانحراف المعيارى للقيم التالية:

20	19	13	48	19	32	22	17	35

الحل:

نكون الجدول التالي:

س	$ح_س$	$ح_س^2$
35	10	100
17	8-	64
22	3-	9
32	7	49
19	6-	36
48	23	529
13	12-	144
19	6-	36
20	5-	25
225	-	992

حساب المتوسط:

$$م_س = \frac{مجـ س}{ن_1} = \frac{225}{9} = 25$$

بعد حساب مس نحسب عمود ح ومنه نحسب ح2 ثم نعوض فى القانون:

$$ع^2 = \frac{مجـ ح^2}{ن}$$

$$ع^2 = \frac{992}{9} = 110.22$$

ومنه نحسب $ع = \sqrt{110.22} = 10.5$

التباين والانحراف المعيارى من البيانات المبوبة:

يحسب التباين من البيانات المبوبة من العلاقة التالية:

$$ع^2 = ل^2 \times \left\{ \frac{مجـ(ح^2 \times ك)}{مجـ ك} - \left\{ \frac{مجـ(ح \times ك)}{مجـ ك} \right\}^2 \right\}$$

حيث:

$ع^2$ = التباين

ل = طول الفئة

ح = الانحراف ويحسب عن طريق وضع صفر فى عموده أمام الفئة التى يقابلها أكبر تكرار ثم نضع من أسفل (1، 2، 3، وهكذا) ومن أعلى نضع (1-، 2-، 3-، وهكذا)

مثال:

احسب الانحراف المعيارى من الجدول التالى:

فئات الدخل	10 -	20 -	30 -	40 -	50 -	60 -	70 - 80
عدد العمال	5	12	22	38	22	12	5

الحل:

نكون الجدول التالى:

ف	ك	ح	ح×ك	$ح^2$×ك
10 -	5	3-	15-	225
20 -	12	2-	24-	576

30 -	22	-1	-22	484
40 -	38	0	0	0
50 -	22	1	22	484
60 -	12	2	24	576
-80 70	5	3	15	225
مج	116	-	صفر	2570

ثم نعوض في القانون:

$$ع^2 = ل^2 \times \left\{ \frac{مجـ(ح^2 \times ك)}{مجـ ك} - \left\{ \frac{مجـ(ح \times ك)}{مجـ ك} \right\}^2 \right\}$$

$$ع^2 = (10)^2 \times \left\{ \frac{2570}{116} - \left\{ \frac{0}{116} \right\}^2 \right\}$$

$$ع^2 = 2215.5$$

$$ع = \sqrt{2215.5} = 47.1$$

ثالثاً: الانحراف المتوسط

الانحراف المتوسط من البيانات الغير مبوبة (المفردة)

$$\text{الانحراف المتوسط} = \frac{\text{مجـ}\,|س - س/|}{ن}$$

حيث:

س = القيمة

س/ = متوسط القيم

ن = عدد القيم

مثال:

لمجموعة البيانات التالية احسب الانحراف المتوسط: -

2 – 3 – 5 – 6 – 7 – 8 – 8 – 9

الحل:

نحسب س/ = (2 + 3 + 5 + 6 + 7 + 8 + 8 + 9) / 8 = 48 / 8 = 6

نكون الجدول التالى:

س	|س – س/|
2	4
3	3

5	1
6	0
7	1
8	2
8	2
9	3
مج	16

$$\text{الانحراف المتوسط} = \frac{16}{8} = 2$$

الانحراف المتوسط من البيانات المبوبة

$$\text{الانحراف المتوسط} = \frac{\text{مجـ}(|\text{س} - \text{س/}| \times \text{ك})}{\text{مجـ ك}}$$

مثال:

من بيانات الجدول التالى احسب الانحراف المتوسط: -

الفئات	– 16	– 20	– 24	– 28	36 – 32
عدد المبحوثين	10	15	40	20	15

الحل:

نكون الجدول التالى:

ف	ك	س	ح/	ح/ × ك	|س - س/|	|س - س/| × ك
16 -	10	18	2-	20-	8.6	86
20 -	15	22	1-	15-	4.6	69
24 -	40	26	0	0	0.6	24
28 -	20	30	1	20	3.4	68
36 32-	15	34	2	30	7.4	111
مج	100	مج		15	مج	358

$$\text{س}/ = 26 + \frac{15}{100} \times 4 = 26 + 0.6 = 26.6$$

$$\text{الانحراف المتوسط} = \frac{358}{100} = 3.58$$

الالتواء وتحديد اعتدالية التوزيع

$$\text{الالتواء} = \frac{3\ (\text{م} - \text{و})}{\text{ع}}$$

حيث:

م: المتوسط

و: الوسيط

ع: الانحراف المعيارى.

مثال:

حدد نوع التوزيع التالى:

10 – 40 – 60 – 50 – 20

الحل:

حساب المتوسط:

$$م = \frac{مجـ س}{ن} = \frac{10 + 40 + 60 + 50 + 20}{5} = 36$$

حساب الوسيط:

نرتب القيم تصاعدياً:

60	50	40	20	10

الوسيط = 40

حساب الانحراف المعيارى:

$$ع = \sqrt{\frac{مجـ ح^2}{ن}}$$

نكون الجدول التالى:

$ح^2$	ح	س
256	16–	20
196	14	50

60	24	576
40	4	16
10	26−	676
مج	0	1720

$$ع = \sqrt{\frac{مج\ ح^2}{ن}} = \sqrt{\frac{1720}{5}} = 18.54$$

$$الالتواء = \frac{3\,(م - و)}{ع} = \frac{3\,(40 - 36)}{18.54} = -\,0.64$$

الالتواء قيمته سالبة فيكون التواء التوزيع سالب.

تمارين

1 – فيما يلى مجموعة بيانات هى:

95 – 200 – 250 – 300 – 110 – 90 – 150 – 100 – 350 – 80

المطلوب حساب:

- المدى
- التباين
- الانحراف المعيارى
- المتوسط
- الوسيط
- المنوال
- الانحراف المتوسط
- حدد نوع الالتواء

2 – لمجموعة البيانات التالية احسب الانحراف المتوسط: –

2 – 3 – 5 – 6 – 7 – 8 – 8 – 9

3 – فيما يلى الدرجات التى حصل عليها عشرة طلاب فى اختبار مادة الإحصاء وهى:

12 – 3 – 5 – 5 – 4 – 10 – 3 – 9 – 10 – 12

المطلوب حساب:

- المدى
- التباين
- الانحراف المعيارى
- المتوسط
- الوسيط
- المنوال
- الانحراف المتوسط
- حدد نوع الالتواء

4 – فيما يلى أعمار 10 طلاب بالفرقة الأولى قسم الاجتماع

17 – 18 – 19 – 17 – 18 – 20 – 22 – 17 – 18 – 19

المطلوب حساب:

- الانحراف المتوسط
- تحديد نوع الالتواء

5 – من بيانات الجدول التالى احسب: –

الفئات	16 –	20 –	24 –	28 –	32 – 36
عدد المبحوثين	10	15	40	20	15

- المدى
- التباين
- الانحراف المعيارى
- المتوسط
- الوسيط
- المنوال
- الانحراف المتوسط
- حدد نوع الالتواء

الفصل السابع

تحليل التباين

مقدمه:

أولاً: طريقة حساب نسبة ف

ثانياً: تحديد مدى دلالة نسبة ف من عدمه.

مقدمـــة

دلت الأبحاث الإحصائية التى قام بها فيشر على أهمية تحليل التباين فى الميادين المختلفة لعلوم الحياة وخاصة فى الكشف عن مدى تجانس العينات ومدى انتسابها إلى أصل واحد أو أصول متعددة.

وبالطبع هناك تساؤل لماذا نستخدم تحليل التباين «النسبة الفائية» للحكم على دلالة الإحصائية للعلاقة بين متغيرين وقد استخدمنا من قبل اختبار «ت» لنفس الغرض.

الإجابة بمنتهى السهولة هو أن اختبار «ت» يستخدم لدراسة العلاقة بين متغيرين فقط لا غير أما إذا زاد عدد المتغيرات عن اثنين فلا يمكن استخدام اختبار «ت» بل نستخدم «نسبة ف».

وبالتالي فان «نسبة ف» تصلح فى حالة متغيرين أو أكثر.

ويعتمد تحليل التباين فى صورته النهائية على قياس مدى اقتراب التباين الداخلي من التباين الخارجي أو مدى ابتعاده عنه وتقاس هذه الناحية بالنسبة التباينية أو النسبة الفائية من خلال العلاقة:

$$\text{نسبة ف} = \frac{\text{التباين الكبير}}{\text{التباين الصغير}}$$

حيث أن التباين الكبير هو الأكبر فى القيمة والتباين الصغير هو الأصغر فى القيمة.

طريقة حساب نسبة ف

- **حساب التباين بين المجموعات**

$$\text{التباين بين المجموعات} = \frac{\text{مجموع المربعات بين المجموعات}}{\text{درجة حرية التباين بين المجموعات}}$$

مجموع المربعات بين المجموعات = $ن_1 ق^2_1 + ن_2 ق^2_2 + ن_3 ق^2_3 +$ 00000000000

حيث:

$ن_1$، $ن_2$، $ن_3$،.... هى عدد أفراد المجموعات

$ق^2_1$، $ق^2_2$، $ق^2_3$،.... هى مربع انحراف متوسط كل مجموعة عن المتوسط الكلى للمجموعات ويحسب من العلاقة:

$ق^2_1 = (م - م_1)^2$

حيث "م" هو المتوسط الوزنى أو الكلى لكافة المجموعات.

درجة حرية التباين بين المجموعات = عدد المجموعات – 1

- **حساب التباين داخل المجموعات**

$$\text{التباين داخل المجموعات} = \frac{\text{مجموع المربعات داخل المجموعات}}{\text{درجة حرية التباين داخل المجموعات}}$$

مجموع المربعات داخل المجموعات = $ن_1 ع^2_1 + ن_2 ع^2_2 + ن_3 ع^2_3 +$ 00000000000

حيث:

$ن_1$، $ن_2$، $ن_3$،.... هى عدد أفراد المجموعات

$ع^2_1$، $ع^2_2$، $ع^2_3$،.... هو التباين لكل مجموعة ويحسب من العلاقة:

$$ع_1^2 = \frac{\text{مجـ س}^2}{\text{ن}} - \left(\frac{\text{مجـ س}}{\text{ن}}\right)^2$$

درجة حرية التباين داخل المجموعات = مجموع أفراد جميع المجموعات – عدد المجموعات

تحديد مدى دلالة "نسبة ف" من عدمه

على أى حال نحصل من قانون "نسبة ف" على "ف" المحسوبة نقوم بمقارنتها بـ "ف" الجدوليه ونتبع الآتى:

إذا كانت قيمة "ف" المحسوبة < قيمة "ف" الجدولية فان "نسبة ف" تكون دالة إحصائية.

أما إذا كانت قيمة "ف" المحسوبة > قيمة "ف" الجدولية فان "نسبة ف" ليست دالة إحصائية.

مثال:

درجات الذكور	23	21	19	19	18
درجات الإناث	19	19	18	14	15

الجدول السابق يوضح درجات خمس ذكور وخمس إناث فى اختبار ما والمطلوب حساب النسبة الفائية وبيان مدى دلالتها إحصائيا عند مستوى دلالة 0.05 وكذلك عند مستوى دلالة 0.01؟

الحل:

نفترض أن درجات الذكور هى "س" ودرجات الإناث هى "ص" ثم نكون الجدول التالى:

س	$س^2$	ص	$ص^2$
23	529	19	361
21	441	19	361
19	361	18	324
19	361	14	196
18	314	15	225
100	2016	85	1467

حساب المتوسطات:

$$م س = \frac{مجـ س}{ن_1} = \frac{100}{5} = 20$$

$$م ص = \frac{مجـ ص}{ن_2} = \frac{85}{5} = 17$$

حساب المتوسط الكلى:

$$م = \frac{م_س + م_ص}{2} = \frac{20 + 17}{2} = 18.5$$

حساب مربع انحراف كل متوسط عن المتوسط الكلى:

$$ق^2_س = (م - م س)^2 = (18.5 - 20)^2 = (-1.5)^2 = 2.25$$

$$ق^2_ص = (م - م ص)^2 = (18.5 - 17)^2 = (1.5)^2 = 2.25$$

حساب التباين:

$$\text{ع}^2_{\text{س}} = \frac{\text{مجـ س}^2}{\text{ن}} - \left(\frac{\text{مجـ س}}{\text{ن}}\right)^2$$

$$\text{ع}^2_{\text{س}} = \frac{2016}{5} - \left(\frac{100}{5}\right)^2 = 3.2$$

$$\text{ع}^2_{\text{ص}} = \frac{\text{مجـ ص}^2}{\text{ن}} - \left(\frac{\text{مجـ ص}}{\text{ن}}\right)^2$$

$$\text{ع}^2_{\text{س}} = \frac{1467}{5} - \left(\frac{85}{5}\right)^2 = 4.4$$

حساب مجموع المربعات بين المجموعات:

مجموع المربعات بين المجموعات = $\text{ن}_1\text{ق}^2_{\text{س}} + \text{ن}_2\text{ق}^2_{\text{ص}}$

مجموع المربعات بين المجموعات = $5 \times 2.25 + 5 \times 2.25$

مجموع المربعات بين المجموعات = 22.5

حساب مجموع المربعات داخل المجموعات:

مجموع المربعات داخل المجموعات = $\text{ن}_1\text{ع}^2_{\text{س}} + \text{ن}_2\text{ع}^2_{\text{ص}}$

مجموع المربعات داخل المجموعات = $5 \times 3.2 + 5 \times 4.4$

مجموع المربعات داخل المجموعات = 38

حساب درجات الحرية:

درجة حرية التباين بين المجموعات = عدد المجموعات – 1

درجة حرية التباين بين المجموعات = 2 – 1 = 1

درجة حرية التباين داخل المجموعات = عدد أفراد جميع المجموعات – عدد المجموعات

درجة حرية التباين داخل المجموعات = 5 + 5 – 2 = 8

حساب التباين بين المجموعات:

$$\text{التباين بين المجموعات} = \frac{\text{مجموع المربعات بين المجموعات}}{\text{درجة حرية التباين بين المجموعات}}$$

$$\text{التباين بين المجموعات} = \frac{22.5}{1} = 22.5$$ **(الأكبر)**

حساب التباين داخل المجموعات:

$$\text{التباين داخل المجموعات} = \frac{\text{مجموع المربعات داخل المجموعات}}{\text{درجة حرية التباين داخل المجموعات}}$$

$$\text{التباين داخل المجموعات} = \frac{38}{8} = 4.75$$ **(الأصغر)**

حساب نسبة ف:

$$\text{نسبة ف} = \frac{\text{التباين الكبير}}{\text{التباين الصغير}}$$

$$\text{نسبة ف} = \frac{22.5}{4.75} = 4.7368$$

حساب "ف" الجدولية:

لحساب "ف" الجدولية نستخدم درجة حرية التباين الكبير = 1 ودرجة حرية التباين الصغير = 8 ونبحث في جداول النسبة الفائية بدرجتي الحرية السابقتين فنحصل على القيمتين:

"ف" الجد ولية = 5.32 عند مستوى دلالة 0.05

"ف" الجد ولية = 11.26 عند مستوى دلالة 0.01

تحديد مدى دلالة "نسبة ف"

- "نسبة ف" المحسوبة = 4.7 < "ف" الجدولية عند مستوى دلالة 0.05 = 5.32، لذا فان "نسبة ف" ليست دالة عند مستوى 0.05.
- "نسبة ف" المحسوبة = 4.7 < "ف" الجدولية عند مستوى دلالة 0.01 = 11.26، لذا فان "نسبة ف" ليست دالة عند مستوى 0.01.

التعليق:

يمكن القول بأن جميع الفروق التي حصل عليها الباحث ليس لها دلالة إحصائية ولا توجد فروق معنوية بين المجموعتين وهذه الفروق ليست إلا مجرد صدفة.

مثال:

س	4	5	7	9	11	–
ص	3	6	8	11	13	22
هـ	7	9	13	16	–	–

الجدول السابق يوضح ثلاث مجموعات من الطلاب فى اختبار ما والمطلوب حساب النسبة الفائية وبيان مدى دلالتها إحصائيا عند مستوى دلالة 0.05 وكذلك عند مستوى دلالة 0.01؟

الحل:

نكون الجدول التالى:

س	ص	هـ	$س^2$	$ص^2$	$هـ^2$
4	3	7	16	9	49
5	6	9	25	36	81
7	8	13	49	64	169
9	11	16	81	121	256
11	13	–	121	169	–
–	22	–	–	484	
36	63	45	292	883	555

حساب المتوسطات:

$$م\ س = \frac{مجـ\ س}{ن_1} = \frac{36}{5} = 7.2$$

$$م\ ص = \frac{مجـ\ ص}{ن_2} = \frac{63}{6} = 10.5$$

$$م\ هـ = \frac{مجـ\ هـ}{ن_2} = \frac{45}{4} = 11.25$$

حساب المتوسط الكلي:

$$م = \frac{م_{س} + م_{ص} + م_{هـ}}{3} = \frac{11.25 + 10.5 + 7.2}{3} = 9.65$$

حساب مربع انحراف كل متوسط عن المتوسط الكلي:

$$ق^2 س = (م - م\ س)^2 = (9.65 - 7.2)^2 = (2.45)^2 = 6$$

$$ق^2 ص = (م - م\ ص)^2 = (9.65 - 10.5)^2 = (-0.85)^2 = 0.7225$$

$$ق^2 هـ = (م - م\ هـ)^2 = (9.65 - 11.25)^2 = (-1.6)^2 = 2.56$$

حساب التباين:

$$ع^2_{س} = \frac{مجـ\ س^2}{ن_1} - \left(\frac{مجـ\ س}{ن_1}\right)^2$$

$$ع^2_{س} = \frac{292}{5} - \left(\frac{36}{5}\right)^2 = 6.56$$

$$ع^2_{ص} = \frac{مجـ\ ص^2}{ن_2} - \left(\frac{مجـ\ ص}{ن_2}\right)^2$$

$$ع^2_{س} = \frac{883}{6} - \left(\frac{63}{6}\right)^2 = 36.92$$

$$ع^2_{هـ} = \frac{مجـ\ هـ^2}{ن_3} - \left(\frac{مجـ\ هـ}{ن_3}\right)^2$$

$$ع^2_{هـ} = \frac{555}{4} - \left(\frac{45}{4}\right)^2 = 12.18$$

حساب مجموع المربعات بين المجموعات:

مجموع المربعات بين المجموعات = $ن_1 ق^2_{س} + ن_2 ق^2_{ص} + ن_3 ق^2_{هـ}$.

مجموع المربعات بين المجموعات = $5 \times 6 + 6 \times 0.7225 + 4 \times 2.56$

مجموع المربعات بين المجموعات = 44.57

حساب مجموع المربعات داخل المجموعات:

مجموع المربعات داخل المجموعات = $ن_1 ع^2_{س} + ن_2 ع^2_{ص} + ن_3 ع^2_{هـ}$

مجموع المربعات داخل المجموعات = $5 \times 6.56 + 6 \times 36.92 + 4 \times 12.18$

مجموع المربعات داخل المجموعات = 303.04

حساب درجات الحرية:

درجة حرية التباين بين المجموعات = عدد المجموعات – 1

درجة حرية التباين بين المجموعات = 3 – 1 = 2

درجة حرية التباين داخل المجموعات = عدد أفراد جميع المجموعات – عدد المجموعات

درجة حرية التباين داخل المجموعات = 5 + 6 + 4 – 3 = 12

حساب التباين بين المجموعات:

$$\text{التباين بين المجموعات} = \frac{\text{مجموع المربعات بين المجموعات}}{\text{درجة حرية التباين بين المجموعات}}$$

$$\text{التباين بين المجموعات} = \frac{44.57}{2} = 22.27$$ **(الأصغر)**

حساب التباين داخل المجموعات:

$$\text{التباين داخل المجموعات} = \frac{\text{مجموع المربعات داخل المجموعات}}{\text{درجة حرية التباين داخل المجموعات}}$$

$$\text{التباين داخل المجموعات} = \frac{303.04}{8} = 24.2 \quad \text{(الأكبر)}$$

حساب نسبة ف:

$$\text{نسبة ف} = \frac{\text{التباين الكبير}}{\text{التباين الصغير}}$$

$$\text{نسبة ف} = \frac{25.2}{22.27} = 1.13$$

حساب "ف" الجدولية:

لحساب "ف" الجدولية نستخدم درجة حرية التباين الكبير = 12 ودرجة حرية التباين الصغير = 2 ونبحث في جداول النسبة الفائية بدرجتي الحرية السابقتين فنحصل على القيمتين:

"ف" الجد ولية = 19.41 عند مستوى دلالة 0.05

"ف" الجد ولية = 99.42 عند مستوى دلالة 0.01

تحديد مدى دلالة "نسبة ف"

- "نسبة ف" المحسوبة = 1.13 < "ف" الجدولية عند مستوى دلالة 0.05 = 19.41، لذا فان "نسبة ف" ليست دالة عند مستوى 0.05
- "نسبة ف" المحسوبة = 1.13 < "ف" الجدولية عند مستوى دلالة 0.01 = 99.42، لذا فان "نسبة ف" ليست دالة عند مستوى 0.01

تمارين

1 - الجدول التالى يوضح درجات مجموعتين من الطلاب فى اختبار فى مادة الإحصاء الاجتماعي:

س	19	20	14	17	25	مجـ = 95
ص	14	11	12	13	20	مجـ = 70

والمطلوب حساب نسبة "ف" مع بيان عما إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

2 - الجدول التالى يوضح درجات مجموعتين من الطلاب فى اختبار فى مادة الحاسب الآلى:

س	18	19	13	16	24	مجـ = 90
ص	15	12	13	14	21	مجـ = 75

والمطلوب حساب نسبة "ف" مع بيان عما إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

3 – الجدول التالى يوضح درجات مجموعتين من الطلاب فى اختبار فى مادة اللغة الفرنسية:

س	7	15	15	11	12
ص	7	2	8	7	6

احسب الدلالة للفروق القائمة بين تلك الدرجات بطريقة تحليل التباين عند مستوى دلالة 0.05

4 – الجدول التالى يوضح درجات 4 مجموعات من الطلاب فى اختبار فى مادة اللغة العربية:

س	49	59	61	60	61
ص	68	55	60	67	60
ع	64	63	54	52	62
هـ	67	55	65	64	59

احسب الدلالة الإحصائية للفروق القائمة بين تلك الدرجات بطريقة تحليل التباين عند مستوى دلالة 0.05 وبين مدى تجانس هذه المجموعات بالنسبة لأصل واحد أو لأصول متعددة.

الفصل الثامن

اختبار «ت»

مقدمه:

أولاً: شروط استخدام اختبار «ت»

ثانياً: الحالات المختلفة لحساب «ت»

مقدمـــة

يعد اختبار «ت» من أكثر اختبارات الدلالة شيوعاً فى الأبحاث النفسية والاجتماعية والتربوية، وترجع نشأته الأولى إلى أبحاث العالم «ستودنت» ولهذا سمى الاختبار بأكثر الحروف تكراراً فى اسمه وهو حرف التاء.

ومن أهم المجالات التى يستخدم فيها هذا الاختبار الكشف عن الفروق بين تحصيل الذكور والإناث فى مادة دراسية ما وذلك عن طريق حساب دلالة فرق متوسط تحصيل الذكور عن متوسط تحصيل الإناث.

ويمكن القول أن اختبار «ت» يستخدم لقياس دلالة فروق المتوسطات غير المرتبطة والمرتبطة للعينات المتساوية والغير متساوية.

شروط استخدام اختبار «ت» لدلالة فروق المتوسطات

لا يحق للباحث أن يستخدم اختبار «ت» قبل أن يدرس خصائص متغيرات البحث من النواحى التالية: -

1 - حجم كل عينة.

2 - الفرق بين حجم عينتي البحث.

3 - مدى تجانس العينة.

4 - مدى اعتدالية التوزيع التكرارى لكل من عينتى البحث.

1 – حجم كل عينة

يجب أن يزيد حجم كل من العينتين عن «5» ويفضل أن يزيد عن «30» أما إذا قل حجم أى من العينتين عن «5» فلا يمكن استخدام اختبار «ت».

2 – الفرق بين حجم عينتى البحث: شرط التقارب

يجب أن يكون حجم عينتى البحث متقارباً فلا يكون مثلاً حجم أحد العينتين «500» وحجم الأخرى «30» لأن للحجم أثره على مستوى دلالة «ت».

3 – مدى تجانس العينتين

يقصد بتجانس العينات مدى انتسابها إلى أصل واحد أو أصول متعددة. فإذا انتسبت العينات إلى أصل واحد فهى متجانسة وإذا لم تنتسب العينات إلى أصل واحد فهى غير متجانسة.

وبالطبع يصعب بالنسبة للباحث تحديد أصول العينات لتحديد تجانسها لذا يمكنه استخدام النسبة الفائية لتحديد التجانس.

يحدد تجانس العينتين من خلال حساب قيمة النسبة الفائية حيث تحسب من العلاقة:

$$\text{ف} = \frac{\text{التباين الأكبر}}{\text{التباين الأصغر}}$$

حيث أن التباين الأكبر هو التباين الأكبر فى القيمة دون التحيز لأحد العينتين، والتباين الأصغر هو الأصغر فى القيمة دون التحيز لأحد العينتين.

بالطبع نحصل من القانون السابق على قيمة لـ «ف» تسمى بقيمة ف المحسوبة ولتحديد التجانس نحسب قيمة أخرى تسمى ف الجدولية ونحصل عليها من جداول «ف» الإحصائية عند درجة حرية التباين الأكبر ودرجة حرية التباين الأصغر ومستوى الدلالة الذى قيمته إما «0.05» أو «0.01» حيث نحسب درجات الحرية من القانون التالى:

درجة حرية التباين الأصغر = ن – 1

حيث «ن» هى عدد أفراد العينة التى تبيانها هو الأكبر.

درجة حرية التباين الأصغر = ن – 1

حيث «ن» هى عدد أفراد العينة التى تبيانها هو الأصغر.

تحديد التجانس

- إذا كانت قيمة «ف» المحسوبة < قيمة "ف" الجدولية فلا يوجد هناك تجانس.
- أما إذا كانت قيمة "ف" المحسوبة > قيمة "ف" الجدولية فيوجد هناك تجانس.

4 – مدى اعتدالية التوزيع التكرارى لكل من العينتين

يكون التوزيع التكرارى معتدلاً عندما تكون قيمة الالتواء الخاص به محصورة بين القيمتين [– 3، + 3] أى واقعة فى الفترة المغلقة – 3 و + 3.

ويحسب الالتواء من القانون التالى: –

$$\text{الالتواء} = \frac{3\,(\text{م} - \text{و})}{\text{ع}}$$

حيث:

- «م» هو المتوسط الحسابى ويحسب من العلاقة

$$\text{م} = \frac{\text{مجـ س}}{\text{ن}}$$

حيث: «مجـ س» هى مجموع القيم، س هى القيم، ن هى عدد القيم.

- «و» هو الوسيط، ويحسب عن طريق ترتيب القيم تصاعدياً أو تنازلياً ثم اختيار قيمة الوسيط فى حالة أن يكون عدد الأفراد فردياً تكون قيمة الوسيط التى ترتيبها (ن+ 1)/ 2أما إذا كان عدد الأفراد زوجياً فتكون قيمة الوسيط هى متوسط القيمتين اللتان ترتيبهما ن/ 2، ن/ 2 + 1.

• «ع» هو الانحراف المعيارى ويحسب من العلاقة:

$$ع^2 = \frac{مجـ ح^2}{ن}$$

من الواضح أن القانون السابق يحسب قيمة التباين فنأخذ للقيمة الناتجة الجذر التربيعى لنحصل على الانحراف المعيارى كالتالى.

$$ع = \sqrt{\frac{مجـ ح^2}{ن}}$$

حيث:

ع = الانحراف المعيارى

ح = الانحراف = س – م

ن = عدد القيم

تحديد مدى دلالة "ت" من عدمه

سنحصل فى جميع حالات «ت» على قيمة لـ «ت» نسميها «ت المحسوبة» ثم نقارنها بقيمة لـ «ت» نحصل عليها من الجداول تسمى «ت الجدولية»

• إذا كانت قيمة «ت المحسوبة» < قيمة "ت الجدولية" تكون قيمة "ت" دالة إحصائية.

• أما إذا كانت قيمة "ت المحسوبة" > قيمة "ت الجدولية" تكون قيمة "ت" ليست دالة إحصائية.

الحالات المختلفة لحساب "ت"

1 – الحالة الأولى: حساب "ت" لدلالة فرق عينتين متجانستين غير متساويتين فى أعداد أفرادهما.

فى هذه الحالة تكون $ن_1$ لا تساوى $ن_2$ حيث $ن_1$، $ن_2$ هما عدد أفراد العينة الأولى والثانية على الترتيب.

تحسب دلالة "ت" لفرق عينتين متجانستين ومختلفين فى عدد الأفراد بالمعادلة التالية

$$ت = \frac{م_1 - م_2}{\sqrt{\left(\frac{ن_1 ع_1^2 + ن_2 ع_2^2}{ن_1 + ن_2 - 2}\right)\left(\frac{1}{ن_1} + \frac{1}{ن_2}\right)}}$$

حيث:

م1 = المتوسط الحسابى للمجموعة الأولى.

م2 = المتوسط الحسابى للمجموعة الثانية.

ع1 2 = تباين المجموعة الأولى.

ع2 2 = تباين المجموعة الثانية.

ن1 = عدد أفراد المجموعة الأولى.

ن2 = عدد أفراد المجموعة الثانية.

مثال:

العينة الأولى	7	4	5	3	8	6	2
العينة الثانية	3	5	15	2	10	13	-

الجدول السابق يوضح درجات مجموعة من الذكور والإناث في اختبار للذكاء والمطلوب حساب قيمة "ت" من خلال التحقق من شروط اختبار "ت" ومن ثم تحديد هل "ت" دالة إحصائية أم لا؟ عند مستوى دلالة إحصائية 0.01؟

الحل:

قبل أن نبدأ الحل نلاحظ أن

$ن_1 = 7 \neq ن_2 = 6$

نعتبر أن العينة الأولى هى "س" والعينة الثانية هى "ص" ونقوم ببناء الجدول التالى.

س	$ح_س$	$ح_س^2$	ص	$ح_ص$	$ح_ص^2$
7	2	4	3	5-	25
4	1-	1	5	3-	9
5	0	0	15	7	49
3	2-	4	2	6-	36
8	3	9	10	4	16
6	1	1	13	5	25
2	3-	9	-	-	-
35	-	28	48	-	148

العينة الأولى:

نحسب لها المتوسط والوسيط والتباين والانحراف المعيارى كالتالى:

حساب المتوسط:

$$م_س = \frac{مجـ\ س}{ن_1} = \frac{35}{7} = 5$$

حساب الوسيط:

نرتب قيم المتغير (س) ترتيباً تصاعدياً كالتالى:

8	7	6	5	4	3	2

حيث أن عدد أفراد العينة الأولى فردية لذا فان قيمة الوسيط هى القيمة التى ترتيبها (ن+ 1 / 2) أى التى ترتيبها (4)

الوسيط = و س = 5

حساب التباين:

$$ع^2 س = \frac{مجـ ح^2 س}{ن_1} = \frac{28}{7} = 4$$

حساب الانحراف المعيارى:

$$ع س = \sqrt{ع^2 س} = \sqrt{4} = 2$$

حساب الالتواء:

$$الالتواء = \frac{3\ (م - و)}{ع} = \frac{3\ (5 - 5)}{2} = صفر$$

العينة الثانية:

نحسب لها المتوسط والوسيط والتباين والانحراف المعيارى كالتالى:

حساب المتوسط:

$$م ص = \frac{مجـ ص}{ن_2} = \frac{48}{6} = 8$$

حساب الوسيط:

نرتب قيم المتغير (ص) ترتيباً تصاعدياً كالتالى:

15	13	10	5	3	2

حيث أن عدد أفراد العينة الثانية زوجية لذا فان قيمة الوسيط هى متوسط القيمتين اللتان ترتيبها (ن/ 2، ن/ 2 + 1) أى التى ترتيبها (3، 4)

الوسيط = وص = (5 + 10)/ 2 = 7.5

حساب التباين:

$$ع^2 ص = \frac{مجـ ح^2 ص}{ن_2} = \frac{148}{6} = 24.66$$

حساب الانحراف المعيارى:

$$ع ص = \sqrt{ع^2 ص} = \sqrt{24.66} = 5$$

حساب الالتواء:

$$الالتواء = \frac{3 (م - و)}{ع} = \frac{3 (8 - 7.5)}{2} = 0.3$$

التحقق من شروط اختبار «ت»

1 - حجم العينتين:

$ن_1 = 7 > 5$

$ن_2 = 6 > 5$

حيث أن حجم كل من العينتين على حده لابد وأن يكون أكبر من 5 لذا فهذا الشرط متحقق.

2 – تقارب العينتين:

ن$_{1}$ = 7 تتقارب جداً من ن$_{2}$ = 6

3 – تجانس العينتين:

نحسب قيمة "ف" المحسوبة من العلاقة:

$$\text{ف المحسوبة} = \frac{\text{التباين الأكبر}}{\text{التباين الأصغر}} = \frac{24.66}{4} = 6.116$$

لإيجاد قيمة "ف" الجدولية يلزم حساب قيمة كل من درجة حرية التباين الأكبر ودرجة حرية التباين الأصغر.

درجة حرية التباين الأكبر = ن$_{2}$ – 1 = 6 – 1 = 5

ونلاحظ أننا اخترنا درجة حرية التباين الأكبر من عدد أفراد المجموعة الثانية لأن تباين العينة الثانية هو الأكبر.

درجة حرية التباين الأصغر = ن$_{1}$ – 1 = 7 – 1 = 6

من جداول "ف" عند درجة حرية تباين كبير (5) ودرجة حرية تباين صغير (6) ومستوى دلالة 0.01 نجد أن قيمة "ف" الجدولية = 8.75.

بمقارنة قيمة "ف" المحسوبة بقيمة "ف" الجدولية نجد أن:

"ف" المحسوبة < "ف" الجدولية (لذا فانه يوجد تجانس بين العينتين).

4 – اعتدالية التوزيع للعينتين:

– 3 < التواء س = صفر < + 3

نلاحظ أن قيمة التواء س محصور فى الفئة] –3،+3 [لذا فان توزيع العينة س معتدل.

– 3 < التواء ص = 0.3 < + 3

نلاحظ أن قيمة التواء ص محصور فى الفئة] –3،+3 [لذا فان توزيع العينة ص معتدل.

حساب قيمة "ت" المحسوبة:

$$ت = \frac{م_1 - م_2}{\sqrt{\left(\frac{ن_1 ع_1^2 + ن_2 ع_2^2}{ن_1 + ن_2 - 2}\right)\left(\frac{1}{ن_1} + \frac{1}{ن_2}\right)}}$$

بالتعويض فى المعادلة السابقة:

$$ت = \frac{5 - 8}{\sqrt{\left(\frac{7 \times 4 + 6 \times 24.66}{7 + 6 - 2}\right)\left(\frac{1}{7} + \frac{1}{6}\right)}}$$

ت المحسوبة = – 1.36

تهمل الإشارة السالبة لقيمة «ت» دائماً فتصبح:

قيمة «ت» المحسوبة = 1.36.

حساب قيمة «ت» الجدولية:

لإيجاد قيمة «ت» الجدولية يلزم حساب درجة الحرية:

درجة الحرية = $ن_1 + ن_2 - 2$ = 7 + 6 – 2 = 11

بالبحث فى جداول «ت» عند درجة حرية 11 ومستوى دلالة 0.01 مع الأخذ فى الاعتبار أن البحث يكون فى دلالة الطرفين، نجد أن قيمة «ت» الجدولية = 3.11.

تحديد دلالة «ت»

بمقارنة قيمة «ت» المحسوبة بقيمة «ت» الجدولية:

نجد أن «ت» المحسوبة = 1.36 < "ت" الجدولية = 3.11

وبالتالي فإن "ت" ليست دالة إحصائية.

2 - الحالة الثانية: حساب "ت" لدلالة فرق عينتين غير متجانستين وغير متساويتين في أعداد أفرادهما

في هذه الحالة تكون $ن_1$ لا تساوى $ن_2$ أيضاً مثل الحالة السابقة حيث $ن_1$، $ن_2$ هما عدد أفراد العينة الأولى والثانية على الترتيب.

تحسب دلالة «ت» لعينتين غير متجانستين ومختلفين في عدد الأفراد بالمعادلة التالية:

$$ت = \frac{م_1 - م_2}{\sqrt{\frac{ع_1^2}{ن_1} + \frac{ع_2^2}{ن_2}}}$$

حيث:

$م_1$ = المتوسط الحسابى للمجموعة الأولى.

$م_2$ = المتوسط الحسابى للمجموعة الثانية.

$ع_1^2$ = تباين المجموعة الأولى.

$ع_2^2$ = تباين المجموعة الثانية.

$ن_1$ = عدد أفراد المجموعة الأولى.

$ن_2$ = عدد أفراد المجموعة الثانية.

مثال:

العينة الأولى	35	17	22	19	32	48	13	19	20
العينة الثانية	11	3	9	10	14	2	7	-	-

الجدول السابق يوضح درجات مجموعة من الذكور والإناث في اختبار للذكاء والمطلوب حساب قيمة "ت" من خلال التحقق من شروط اختبار "ت" ومن ثم تحديد هل "ت" دالة إحصائية أم لا؟ عند مستوى دلالة إحصائية 0.05؟

الحل:

قبل أن نبدأ الحل نلاحظ أن

$ن_1 = 9 \neq ن_2 = 7$

نعتبر أن العينة الأولى هى "س" والعينة الثانية هى "ص" ونقوم ببناء الجدول التالى.

س	$ح_س$	$ح^2_س$	ص	$ح_ص$	$ح^2_ص$
35	10	100	11	3	9
17	8-	64	3	5-	25
22	3-	9	9	1	1
19	6-	36	14	6	36
48	23	569	2	6-	36
13	12-	144	7	1-	1
19	6-	36	-	-	-
20	5-	25	-	-	-
225	-	992	56	-	112

العينة الأولى:

نحسب لها المتوسط والوسيط والتباين والانحراف المعيارى كالتالى:

حساب المتوسط:

$$\text{م س} = \frac{\text{مجـ س}}{\text{ن}_1} = \frac{225}{9} = 25$$

حساب الوسيط:

نرتب قيم المتغير (س) ترتيباً تصاعدياً كالتالى:

13	17	19	19	20	22	32	35	48

حيث أن عدد أفراد العينة الأولى فردية لذا فان قيمة الوسيط هى القيمة التى ترتيبها (ن+ 1 / 2) أى التى ترتيبها (5)

الوسيط = وس = 20

حساب التباين:

$$\text{ع}^2\text{س} = \frac{\text{مجـ ح}^2\text{س}}{\text{ن}_2} = \frac{992}{9} = 110.2$$

حساب الانحراف المعيارى:

$$\text{ع س} = \sqrt{\text{ع}^2\text{س}} = \sqrt{110.2} = 10.5$$

حساب الالتواء:

$$\text{الالتواء} = \frac{3\,(\text{م} - \text{و})}{\text{ع}} = \frac{3\,(25 - 20)}{2} = 1.4$$

العينة الثانية:

نحسب لها المتوسط والوسيط والتباين والانحراف المعيارى كالتالى:

حساب المتوسط:

$$\text{م ص} = \frac{\text{مجـ ص}}{\text{ن}_2} = \frac{56}{7} = 8$$

حساب الوسيط:

نرتب قيم المتغير (ص) ترتيباً تصاعدياً كالتالى:

2	3	7	9	10	11	14

حيث أن عدد أفراد العينة الثانية فردية لذا فان قيمة الوسيط هى القيمة التى ترتيبها (ن+ 1 / 2) أى التى ترتيبها (4)

الوسيط = وص = 9

حساب التباين:

$$\text{ع}^2\text{ص} = \frac{\text{مجـ ح}^2\text{ص}}{\text{ن}_2} = \frac{112}{7} = 16$$

حساب الانحراف المعيارى:

$$\text{ع ص} = \sqrt{\text{ع}^2\text{ص}} = \sqrt{16} = 4$$

حساب الالتواء:

$$\text{الالتواء} = \frac{3\,(\text{م} - \text{و})}{\text{ع}} = \frac{3\,(8 - 9)}{4} = -0.75$$

التحقق من شروط اختبار «ت»

1 – حجم العينتين:

ن1 = 9 > 5

ن2 = 7 > 5

حيث أن حجم كل من العينتين على حده لابد وأن يكون أكبر من 5 لذا فهذا الشرط متحقق.

2 – تقارب العينتين:

$ن_1$ = 9 تتقارب جداً من $ن_2$ = 7

3 – تجانس العينتين:

نحسب قيمة "ف" المحسوبة من العلاقة:

$$\text{ف المحسوبة} = \frac{\text{التباين الأكبر}}{\text{التباين الأصغر}} = \frac{110.2}{16} = 6.88$$

لإيجاد قيمة "ف" الجدولية يلزم حساب قيمة كل من درجة حرية التباين الأكبر ودرجة حرية التباين الأصغر.

درجة حرية التباين الأكبر = $ن_1$ – 1 = 9 – 1 = 8

ونلاحظ أننا اخترنا درجة حرية التباين الأكبر من عدد أفراد المجموعة الأولى لأن تباين العينة الأولى هو الأكبر.

درجة حرية التباين الأصغر = $ن_2$ – 1 = 7 – 1 = 6

من جداول "ف" عند درجة حرية تباين كبير (8) ودرجة حرية تباين صغير (6) ومستوى دلالة 0.05 نجد أن قيمة "ف" الجدولية = 4.15.

بمقارنة قيمة "ف" المحسوبة بقيمة "ف" الجدولية نجد أن:

"ف" المحسوبة > "ف" الجدولية (لذا فانه لا يوجد تجانس بين العينتين).

4 – اعتدالية التوزيع للعينتين:

– 3 < التواء س = 1.4 < + 3

نلاحظ أن قيمة التواء س محصور فى الفئة] 3–،3+ [لذا فان توزيع العينة س معتدل.

– 3 < التواء ص = – 0.75 < + 3

نلاحظ أن قيمة التواء ص محصور فى الفئة] 3–،3+ [لذا فان توزيع العينة ص معتدل.

حساب قيمة "ت" المحسوبة:

$$ت = \frac{م_1 - م_2}{\sqrt{\frac{ع_1^2}{ن_1} + \frac{ع_2^2}{ن_2}}}$$

بالتعويض فى المعادلة السابقة:

$$ت = \frac{25 - 8}{\sqrt{\frac{110.2}{9} + \frac{16}{7}}}$$

ت المحسوبة = 4.46

حساب قيمة «ت» الجدولية:

لإيجاد قيمة «ت» الجدولية

تحسب من العلاقة التالية:

$$ت\ الجدولية = \frac{ت_1 \times (ع_1^2 / ن_1) + ت^2 \times (ع_2^2 / ن_2)}{(ع_1^2 / ن_1) + (ع_2^2 / ن_2)}$$

حيث:

ت$_1$: هى «ت» الجدولية للعينة الأولى وتحسب عن طريق حساب درجة حرية العينة الأولى على حده من العلاقة:

درجة حرية العينة الأولى = ن$_1$ – 1 = 9 – 1 = 8

وبالبحث فى جداول «ت» عن درجة حرية 8 ومستوى دلالة 0.05 فى دلالة الطرفين نجد أن قيمة ت$_1$ = 2.31

ت2: هى «ت» الجدولية للعينة الثانية وتحسب عن طريق حساب درجة حرية العينة الثانية على حده من العلاقة:

درجة حرية العينة الثانية = ن$_2$ – 1 = 7 – 1 = 6

وبالبحث فى جداول «ت» عن درجة حرية 6 ومستوى دلالة 0.05 فى دلالة الطرفين نجد أن قيمة ت$_2$ = 2.45

ثم نعوض فى المعادلة التالية لحساب قيمة «ت» الجدولية:

$$\text{ت الجدولية} = \frac{\text{ت}_1 \times (\text{ع}_1^2/\text{ن}_1) + \text{ت}_2 \times (\text{ع}_2^2/\text{ن}_2)}{(\text{ع}_1^2/\text{ن}_1) + (\text{ع}_2^2/\text{ن}_2)}$$

$$\text{ت الجدولية} = \frac{(7\ /16)\times 2.45 + (9\ /110.2)\times 2.31}{(7\ /16) + (9\ /110.2)}$$

ت الجدولية = 2.33

تحديد دلالة «ت»

بمقارنة قيمة «ت» المحسوبة بقيمة «ت» الجدولية

نجد أن «ت» المحسوبة = 4.46 > "ت" الجدولية = 2.33

وبالتالي فان "ت" دالة إحصائية.

3 - الحالة الثالثة: حساب "ت" لدلالة فرق عينتين غير مرتبطتين ومتساويتين فى أعداد أفرادهما

فى هذه الحالة لا تتحقق من شروط اختبار «ت».

فى هذه الحالة تكون $ن_1 = ن_2$ حيث $ن_1$، $ن_2$ هما عدد أفراد العينة الأولى والثانية على الترتيب.

تحسب دلالة «ت» لفرق عينتين متساويتين فى عدد الأفراد بالمعادلة التالية:

$$ت = \frac{م_1 - م_2}{\sqrt{\frac{ع_1^2 + ع_2^2}{ن - 1}}}$$

حيث:

$م_1$ = المتوسط الحسابى للمجموعة الأولى.

$م_2$ = المتوسط الحسابى للمجموعة الثانية.

$ع_1^2$ = تباين المجموعة الأولى.

$ع_2^2$ = تباين المجموعة الثانية.

ن = عدد أفراد العينة الأولى أو الثانية حيث أنهما متساويتان.

مثال:

العينة الأولى	7	4	5	3	8	6	2
العينة الثانية	3	5	15	2	10	13	1

الجدول السابق يوضح درجات مجموعة من الذكور والإناث في اختبار للذكاء والمطلوب حساب قيمة "ت" ومن ثم تحديد هل "ت" دالة إحصائية أم لا؟ عند مستوى

دلالة إحصائية 0.05؟

الحل:

قبل أن نبدأ الحل نلاحظ أن

$ن_1 = ن_2 = 7$

نعتبر أن العينة الأولى هى "س" والعينة الثانية هى "ص" ونقوم ببناء الجدول التالى.

س	$ح_س$	$ح^2_س$	ص	$ح_ص$	$ح^2_ص$
7	2	4	3	4-	16
4	1-	1	5	2-	4
5	0	0	15	8	64
3	2-	4	2	5-	25
8	3	9	10	3	9
6	1	1	13	6	36
2	3-	9	1	6-	36
35	-	28	49	-	190

العينة الأولى:

نحسب لها المتوسط والتباين.

حساب المتوسط:

$$م_س = \frac{مجـ س}{ن_1} = \frac{35}{7} = 5$$

حساب التباين:

$$ع^2 س = \frac{مجـ ح^2 س}{ن_2} = \frac{28}{7} = 4$$

العينة الثانية:

نحسب لها المتوسط والتباين كالتالى:

حساب المتوسط:

$$م ص = \frac{مجـ ص}{ن_2} = \frac{49}{7} = 7$$

حساب التباين:

$$ع^2 ص = \frac{مجـ ح^2 ص}{ن_2} = \frac{190}{7} = 27.14$$

حساب قيمة "ت" المحسوبة:

$$ت = \frac{م_1 - م_2}{\sqrt{\frac{ع_1^2 + ع_2^2}{ن - 1}}}$$

بالتعويض فى المعادلة السابقة:

$$ت = \frac{7 - 5}{\sqrt{\frac{27.14 + 4}{7 - 1}}}$$

ت المحسوبة = − 0.88

تهمل الإشارة السالبة لقيمة «ت» دائماً فتصبح:

قيمة «ت» المحسوبة = 0.88.

حساب قيمة «ت» الجدولية:

لإيجاد قيمة «ت» الجدولية يلزم حساب درجة الحرية:

درجة الحرية = $ن_1 + ن_2 - 2 = 7 + 7 - 2 = 12$

بالبحث فى جداول «ت» عند درجة حرية 12 ومستوى دلالة 0.05 مع الأخذ فى الاعتبار أن البحث يكون فى دلالة الطرفين، نجد أن قيمة «ت» الجدولية = 2.18.

تحديد دلالة «ت»

بمقارنة قيمة «ت» المحسوبة بقيمة «ت» الجدولية

نجد أن «ت» المحسوبة = 0.88 < "ت" الجدولية = 2.18

وبالتالي فان "ت" ليست دالة إحصائية.

4 – الحالة الرابعة: حساب "ت" لدلالة فرق عينتين مرتبطتين ومتساويتين فى أعداد أفرادهما

يرتبط المتوسطان عندما نجرى اختباراً على مجموعة من الأفراد ثم نعيد نفس الاختبار على نفس المجموعة فى وقت آخر أى أن العينة التى يجرى عليها الاختبار الأول هى نفسها العينة التى يجرى عليها الاختبار الثانى وفى هذه الحالة لا تكون ن1 = ن2 بل تصبح هى نفسها.

فى هذه الحالة أيضاً لا نتحقق من شروط اختبار «ت».

تحسب دلالة «ت» لفرق عينتين متساويتين فى عدد الأفراد بالمعادلة التالية:

$$ت = \frac{م_ف}{\sqrt{\frac{مجـ ح^2 ف}{ن (ن - 1)}}}$$

حيث:

- م ف = متوسط الفروق ويحسب من العلاقة:

$$م_ف = \frac{مجـ ف}{ن}$$

- ف = الفروق = $س_1 - س_2$
- $س_1$ هى درجات الاختبار الأول
- $س_2$ هى درجات الاختبار الثانى
- ن = عدد الأفراد فى أى من الاختبارين.
- $ح_ف = ف - م_ف$

مثال:

درجات الاختبار الأول	26	18	20	24	22	14	23	16	22	11
درجات الاختبار الثانى	23	16	19	21	18	12	24	11	23	9

الجدول السابق يوضح درجات مجموعة من الأطفال في اختبار للذكاء حيث تم إجراء الاختبار مرة ثم بعد إجراء برنامج تدريبى لهم تم إجراء الاختبار مرة أخرى والمطلوب حساب قيمة "ت" للفرق بين درجات الاختبارين ومن ثم تحديد هل "ت" دالة إحصائية أم لا؟ عند مستوى دلالة إحصائية 0.05؟

الحل:

قبل أن نبدأ الحل نلاحظ أن

$ن_1$ هى نفسها $ن_2$

نعتبر أن درجات الاختبار الأول هى "س1" ودرجات الاختبار الثانى هى "س2"

ثم نقوم ببناء الجدول التالى:

$س_1$	$س_2$	ف	$ح_ف$	$ح^2_ف$
26	23	3	1	1
18	16	2	0	0
20	19	1	1-	1
24	21	3	1	1
22	18	4	2	4
14	12	2	0	0
23	24	1-	3-	9
16	11	5	3	9
22	23	1-	3-	9
11	9	2	0	0
–	–	20	–	34

حساب متوسط الفروق م ف:

$$م_ف = \frac{مجـ ف}{ن} = \frac{20}{10} = 2$$

حساب ح ف:

يحسب من العلاقة:

$ح_ف = ف - م_ف$

حساب قيمة "ت" المحسوبة:

$$ت = \frac{م_ف}{\sqrt{\frac{مجـ ح^2 ف}{ن(ن-1)}}}$$

بالتعويض في المعادلة السابقة:

$$ت = \frac{2}{\sqrt{\frac{34}{10(10-1)}}}$$

ت المحسوبة = 3.25

حساب قيمة «ت» الجدولية:

لإيجاد قيمة «ت» الجدولية يلزم حساب درجة الحرية:

درجة الحرية = ن – 1 = 10 – 1 = 9

بالبحث في جداول «ت» عند درجة حرية 9 ومستوى دلالة 0.05 مع الأخذ في الاعتبار أن البحث يكون في دلالة الطرف الواحد، نجد أن قيمة «ت» الجدولية = 1.83.

تحديد دلالة «ت»

بمقارنة قيمة «ت» المحسوبة بقيمة «ت» الجدولية

نجد أن «ت» المحسوبة = 3.25 > "ت" الجدولية = 1.83

وبالتالي فان "ت" دالة إحصائية.

تمارين

1 – القيم التالية تعبر عن الدرجات التي حصل عليها مجموعتين من الأفراد فى مقياس توهم المرض:

س	10	6	9	7	8	5	18
ص	3	5	11	6	8	9	–

حيث س هى مجموعة الذكور، ص مجموعة الإناث.

والمطلوب:

حساب قيمة "ت" بالطريقة المناسبة من خلال التحقق من شروط "ت" مع بيان إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

2 – القيم التالية تعبر عن الدرجات التي حصل عليها مجموعتين من الأفراد فى اختبار للذكاء.

س	27	31	38	25	36	21	39
ص	15	23	27	21	30	19	–

حيث س هى مجموعة الذكور، ص مجموعة الإناث.

والمطلوب:

حساب قيمة "ت" بالطريقة المناسبة من خلال التحقق من شروط "ت" مع بيان إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

3 – القيم التالية تعبر عن الدرجات التي حصل عليها مجموعتين من الأفراد في اختبار يقيس القدرة على التركيز.

س	9	5	8	6	7	4	17
ص	2	4	10	5	7	8	–

حيث س هي مجموعة الذكور، ص مجموعة الإناث.

والمطلوب:

حساب قيمة "ت" بالطريقة المناسبة من خلال التحقق من شروط "ت" مع بيان إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

4 – القيم التالية تعبر عن الدرجات التي حصل عليها مجموعتين من الأفراد في اختبار مادة الإحصاء.

س	8	4	5	3	2	7	9	10
ص	6	5	1	4	9	17	–	–

حيث س هي مجموعة الذكور، ص مجموعة الإناث.

والمطلوب:

حساب قيمة "ت" بالطريقة المناسبة من خلال التحقق من شروط "ت" مع بيان إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

5 – القيم التالية تعبر عن الدرجات التي حصل عليها مجموعتين من الأفراد فى اختبار مادة الحاسب الآلى.

س	9	13	14	12	17
ص	8	3	9	8	17

حيث س هى مجموعة الذكور، ص مجموعة الإناث.

والمطلوب:

حساب قيمة "ت" بالطريقة المناسبة مع بيان إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

6 – القيم التالية تعبر عن الدرجات التي حصل عليها مجموعتين من الأفراد فى اختبار للاستيعاب.

س	7	15	15	11	12
ص	7	2	8	7	6

حيث س هى مجموعة الذكور، ص مجموعة الإناث.

والمطلوب:

حساب قيمة "ت" بالطريقة المناسبة مع بيان إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

7 – قمت بتطبيق اختبار على عينة قوامها 10 من الطلاب ثم تم تدريبهم على طريقة الاختبار لمدة أسبوعين وتم إجراء الاختبار مرة أخرى والجدول التالى يوضح درجات الطلاب فى الاختبارين:

درجات الاختبار الأول	30	27	25	32	15	22	17	18	16	28
درجات الاختبار الثانى	25	18	28	16	24	9	26	12	10	14

والمطلوب:

حساب قيمة "ت" بالطريقة المناسبة مع بيان إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

8 – القيم التالية تعبر عن الدرجات التي حصل عليها مجموعتين من الأفراد فى اختبار للذكاء.

س	35	17	22	32	19	48	13	19	20
ص	11	3	9	10	14	2	7	–	–

حيث س هى مجموعة الذكور، ص مجموعة الإناث.

والمطلوب:

حساب قيمة "ت" بالطريقة المناسبة مع بيان إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

الفصل التاسع

اختبار كا^2

مقدمه:

أولاً: الطريقة العامة لحساب كا^2

ثانياً: تحديد مدى دلالة كا^2 من عدمه.

ثالثاً: الطريقة العامة لحساب كا^2 من الجدول التكرارى 1×2.

رابعاً: الطريقة المختصرة لحساب كا^2 من الجدول التكرارى 1×2.

خامساً: الطريقة العامة لحساب كا^2 من الجدول التكرارى 1×ن.

سادساً: الطريقة العامة لحساب كا^2 من الجدول التكرارى 2×2.

سابعاً: الطريقة المختصرة لحساب كا^2 من الجدول التكرارى 2×2.

ثامناً: الطريقة العامة لحساب كا^2 من الجدول التكرارى ن×ن.

تاسعاً: حساب كا^2 لدلالة فروق النسب المرتبطة.

مقدمـــة

ترجع النشأة الأولى لاختبار كا^2 إلى البحث الذي نشره كارل بيرسون فى أوائل القرن العشرين وهى تعد من أهم اختبارات الدلالة الإحصائية وأكثرها شيوعاً لأنها لا تعتمد على شكل التوزيع ولذا فهى تعد من المقاييس اللابارامترية أى مقاييس التوزيعات الحرة ولأنها تحسب لكل خلية من خلايا أى جدول تكرارى ثم تجميع القيم الجزئية للحصول على القيمة الكارلية لـ كا^2.

وتستخدم كا2 لحساب دلالة فروق التكرار أو البيانات العددية التي يمكن تحويلها إلى تكرار مثل النسب والاحتمال.

الطريقة العامة لحساب كا^2

$$\text{كا}^2 = \text{مجـ} \frac{(\text{ت}_\text{و} - \text{ت}_\text{م})^2}{\text{ت}_\text{و}}$$

حيث:

ت$_\text{و}$: هو التكرار الواقعى الذي يحدث بالفعل والموجود بالجدول.

ت$_\text{م}$: هو التكرار المتوقع حدوثه ويختلف حسابه باختلاف نوع الجدول المطلوب حساب كا^2 منه.

تحديد مدى دلالة كا^2 من عدمه

فى جميع الحالات نخرج من الحسابات بقيمة كا2 المحسوبة نقارنها بقيمة كا2 الجدولية كالتالى:

· إذا كانت كا^2 المحسوبة > كا^2 الجدولية فان كا^2 تكون دالة إحصائية.

· إذا كانت كا^2 المحسوبة < كا^2 الجدولية فان كا^2 ليست دالة إحصائية.

حالات حساب كا^2 من الجداول المختلفة:

1 - الحالة الأولى: الطريقة العامة لحساب كا^2 من الجدول التكراري 1×2:

يتكون الجدول 1×2 من صف واحد وعمودين دون خلايا المجموع إن وجدت بالجدول.

ولحساب قيمة كا^2 فى هذا الجدول تحسب من القانون العام:

$$\text{كا}^2 = \text{مجـ} \frac{(\text{ت}_{\text{و}} - \text{ت}_{\text{م}})^2}{\text{ت}_{\text{و}}}$$

حيث تم هنا تساوى متوسط التكرارات الواقعية الموجودة بالجدول.

مثال:

الجدول التالي يوضح آراء 80 شخص فى استبيان دار حول رفض أو قبول قضية الزواج العرفي.

الرأي	موافق	غير موافق	مج
التكرار	60	20	80

والمطلوب حساب قيمة كا^2 مع بيان مدى دلالتها إحصائيا عند مستوى دلالة 0.05؟

الحل:

حساب التكرار المتوقع ($\text{ت}_{\text{م}}$):

$$\text{ت}_{\text{م}} = \frac{20 + 60}{2} = 40$$

حساب كا^2 المحسوبة:

نكون الجدول التالي:

$ت_و$	$ت_م$	$ت_و - ت_م$	$(ت_و - ت_م)^2$	$\frac{(ت_و - ت_م)^2}{ت_م}$
6	40	20	400	10
20	40	20–	400	10
–	–	–	**مجموع**	20

من الجدول مباشرة فان مجموع العمود الأخير يعطينا قيمة كا^2

كا^2 المحسوبة = 20.

حساب كا^2 الجدولية:

لحسابها يلزم حساب درجة الحرية ومستوى الدلالة:

درجة الحرية = عدد الأعمدة – 1 = 2 – 1 = 1

مستوى الدلالة = 0.05.

بالبحث في جداول كا^2 عند درجة حرية = 1 ومستوى دلالة 0.05 نجد قيمة كا^2 الجدولية = 3.841.

تحديد مدى دلالة كا^2:

نقارن قيمة كا^2 المحسوبة بقيمة كا^2 الجدولية نجد أن:

قيمة كا^2 المحسوبة = 20 > قيمة كا^2 الجدولية = 3.841

لذا فان كا^2 دالة إحصائية عند مستوى دلالة 0.05.

2 – الحالة الثانية: الطريقة المختصرة لحساب كا^2 من الجدول التكراري 1×2:

لحساب قيمة كا^2 فى هذا الجدول بالطريقة المختصرة فان قيمة كا^2 من العلاقة:

$$كا^2 = \frac{(ت_1 - ت_2)^2}{ت_1 + ت_2}$$

حيث $ت_1$ هو التكرار الأكبر و $ت_2$ هي التكرار الأصغر.

مثال:

الجدول التالي يوضح آراء 80 شخص فى استبيان دار حول رفض أو قبول قضية الزواج العرفي.

الرأي	موافق	غير موافق	مج
التكرار	60	20	80

والمطلوب حساب قيمة كا2 بالطريقة المختصرة مع بيان مدى دلالتها إحصائيا عند مستوى دلالة 0.05؟

الحل:

حساب كا^2 المحسوبة:

$$كا^2 = \frac{(60 - 20)^2}{60 + 20} = \frac{1600}{80} = 20$$

حساب كا^2 الجدولية:

لحسابها يلزم حساب درجة الحرية ومستوى الدلالة:

درجة الحرية = عدد الأعمدة – 1 = 2 – 1 = 1

مستوى الدلالة = 0.05.

بالبحث فى جداول كا^2 عند درجة حرية = 1 ومستوى دلالة 0.05 نجد قيمة كا^2 الجدولية = 3.841.

تحديد مدى دلالة كا^2:

نقارن قيمة كا^2 المحسوبة بقيمة كا^2 الجدولية نجد أن

قيمة كا^2 المحسوبة = 20 > قيمة كا^2 الجدولية = 3.841

لذا فان كا^2 دالة إحصائية عند مستوى دلالة 0.05.

3 – الحالة الثالثة: الطريقة العامة لحساب كا^2 من الجدول التكراري 1×ن:

يتكون الجدول 1×ن من صف واحد وعدد (ن) عمود دون خلايا المجموع إن وجدت بالجدول.

ولحساب قيمة كا^2 فى هذا الجدول تحسب من القانون العام:

$$\text{كا}^2 = \text{مجـ} \frac{(\text{ت}_{\text{و}} - \text{ت}_{\text{م}})^2}{\text{ت}_{\text{و}}}$$

حيث تم هنا تساوى متوسط التكرارات الواقعية الموجودة بالجدول.

مثال:

الجدول التالي يوضح آراء 30 شخص فى استبيان دار حول قضية الزواج العرفي.

الرأي	موافق	لا أدرى	معارض	مج
التكرار	12	2	16	30

والمطلوب حساب قيمة كا^2 مع بيان مدى دلالتها إحصائيا عند مستوى دلالة 0.05؟

الحل:

حساب التكرار المتوقع ($ت_م$):

$$ت_م = \frac{16 + 2 + 12}{3} = 10$$

حساب كا² المحسوبة:

نكون الجدول التالي:

$ت_و$	$ت_م$	$ت_و - ت_م$	$(ت_و - ت_م)^2$	$\frac{(ت_و - ت_م)^2}{ت_م}$
12	10	2	4	0.4
2	10	8−	64	6.4
16	10	6	36	3.6
−	−	−	**مجموع**	10.4

من الجدول مباشرة فان مجموع العمود الأخير يعطينا قيمة كا²

كا² المحسوبة = 10.4.

حساب كا² الجدولية:

لحسابها يلزم حساب درجة الحرية ومستوى الدلالة:

درجة الحرية = عدد الأعمدة − 1 = 3 − 1 = 2

مستوى الدلالة = 0.05.

بالبحث فى جداول كا2 عند درجة حرية = 2 ومستوى دلالة 0.05 نجد قيمة كا2 الجدولية = 5.991.

تحديد مدى دلالة كا²:

نقارن قيمة كا² المحسوبة بقيمة كا² الجدولية نجد أن

قيمة كا² المحسوبة = 10.4 > قيمة كا² الجدولية = 5.991

لذا فان كا² دالة إحصائية عند مستوى دلالة 0.05.

4 – الحالة الرابعة: الطريقة العامة لحساب كا² من الجدول التكراري 2×2:

يتكون الجدول 2×2 من صفين وعمودين دون خلايا المجموع إن وجدت بالجدول.

ولحساب قيمة كا² فى هذا الجدول تحسب من القانون العام:

$$\text{كا}^2 = \text{مجـ} \frac{(\text{ت}_{\text{و}} - \text{ت}_{\text{م}})^2}{\text{ت}_{\text{و}}}$$

وتحسب تم لكل خلية فى هذا الجدول على حده من العلاقة:

$$\text{ت}_{\text{م}} = \frac{\text{مجموع الصف} \times \text{مجموع العمود}}{\text{المجموع الكلى}}$$

مثال:

الجدول التالي يوضح العلاقة بين المتغيرين النوع وتأييد برنامج تليفزيوني معين.

النوع / الفكرة	ذكور	إناث	المجموع
مؤيد	35	37	72
معارض	14	34	48
المجموع	49	71	120

والمطلوب حساب قيمة كا2 مع بيان مدى دلالتها إحصائيا عند مستوى دلالة 0.05؟

الحل:

حساب التكرار المتوقع ($ت_{م}$):

$ت_{م}$ للخلية الأولى (35)= $\frac{49 \times 72}{120} = 29.4$

$ت_{م}$ للخلية الثانية (37)= $\frac{71 \times 72}{120} = 42.6$

$ت_{م}$ للخلية الثالثة (14)= $\frac{49 \times 48}{120} = 19.6$

$ت_{م}$ للخلية الرابعة (34)= $\frac{71 \times 48}{120} = 28.4$

حساب $كا^2$ المحسوبة:

نكون الجدول التالي:

$ت_{و}$	$ت_{م}$	$ت_{و} - ت_{م}$	$(ت_{و} - ت_{م})^2$	$\frac{(ت_{و} - ت_{م})^2}{ت_{م}}$
35	29.4	5.6	31.36	1.06
37	42.6	5.6−	31.6	0.74
14	19.6	5.6−	31.36	1.6
34	28.4	5.6	31.36	1.1
−	−	−	مجموع	4.5

من الجدول مباشرة فان مجموع العمود الأخير يعطينا قيمة كا^2

كا^2 المحسوبة = 4.5.

حساب كا^2 الجدولية:

لحسابها يلزم حساب درجة الحرية ومستوى الدلالة:

درجة الحرية = (عدد الصفوف – 1) × (عدد الأعمدة – 1)

= (2 – 1) × (2 – 1) = 1 × 1 = 1

مستوى الدلالة = 0.05.

بالبحث في جداول كا^2 عند درجة حرية = 1 ومستوى دلالة 0.05 نجد قيمة كا^2 الجدولية = 3.841

تحديد مدى دلالة كا^2:

نقارن قيمة كا^2 المحسوبة بقيمة كا^2 الجدولية نجد أن:

قيمة كا^2 المحسوبة = 4.5 > قيمة كا^2 الجدولية = 3.841

لذا فان كا^2 دالة إحصائية عند مستوى دلالة 0.05.

5 – الحالة الخامسة: الطريقة المختصرة لحساب كا^2 من الجدول التكراري 2×2:

يتكون الجدول 2×2 من صفين وعمودين دون خلايا المجموع إن وجدت بالجدول.

ولحساب قيمة كا^2 في هذا الجدول بالطريقة المختصرة نطبق القانون التالي:

كا^2 = فاي^2 × ن

حيث:

فاي: هو معامل ارتباط فاي والذى يحسب من العلاقة:

$$\text{فاي} = \frac{\text{أ} \times \text{د} - \text{ب} \times \text{جـ}}{\sqrt{\text{هـ} \times \text{و} \times \text{ز} \times \text{ح}}}$$

حيث أ، ب، جـ، د، هـ، و، ز، ح، ن

هم خلايا الجدول الرباعى الخلايا كما بالشكل التالى:

النوع / الفكرة	ذكور	إناث	المجموع
مؤيد	أ	ب	ح
معارض	جـ	د	ز
المجموع	هـ	و	ن

مثال:

الجدول التالي يوضح العلاقة بين المتغيرين النوع وتأييد برنامج تليفزيوني معين.

النوع / الفكرة	ذكور	إناث	المجموع
مؤيد	35	37	72
معارض	14	34	48
المجموع	49	71	120

والمطلوب حساب قيمة $كا^2$ مع بيان مدى دلالتها إحصائيا عند مستوى دلالة 0.05؟

الحل:

حساب معامل فاى: نعوض فى العلاقة:

$$\text{فاى} = \frac{\text{أ} \times \text{د} - \text{ب} \times \text{جـ}}{\sqrt{\text{هـ} \times \text{و} \times \text{ز} \times \text{ح}}}$$

$$\text{فاى} = \frac{14 \times 37 - 34 \times 35}{\sqrt{72 \times 48 \times 71 \times 49}}$$

فاى = 0.19

حساب كا^2:

كا^2 = فاى^2 × ن

كا^2 = $(0.19)^2 \times 120 = 4.33$

حساب كا^2 الجدولية:

لحسابها يلزم حساب درجة الحرية ومستوى الدلالة:

درجة الحرية = (عدد الصفوف − 1) × (عدد الأعمدة − 1)

$= (2 - 1) \times (2 - 1) = 1 \times 1 = 1$

مستوى الدلالة = 0.05.

بالبحث فى جداول كا^2 عند درجة حرية = 1 ومستوى دلالة 0.05 نجد قيمة كا^2 الجدولية = 3.841

تحديد مدى دلالة كا^2:

نقارن قيمة كا^2 المحسوبة بقيمة كا^2 الجدولية نجد أن:

قيمة كا^2 المحسوبة = 4.33 > قيمة كا^2 الجدولية = 3.841

لذا فان كا^2 دالة إحصائية عند مستوى دلالة 0.05.

6 – الحالة السادسة: الطريقة العامة لحساب كا^2 من الجدول التكراري ن×ن:

يتكون الجدول ن×ن من عدد (ن) من الصفوف وعدد (ن) من الأعمدة دون خلايا المجموع إن وجدت بالجدول.

ولحساب قيمة كا^2 فى هذا الجدول تحسب من القانون العام:

$$كا^2 = مجـ \frac{(ت_و - ت_م)^2}{ت_و}$$

وتحسب تم لكل خلية فى هذا الجدول على حده من العلاقة:

$$ت_م = \frac{\text{مجموع الصف} \times \text{مجموع العمود}}{\text{المجموع الكلى}}$$

مثال:

الجدول التالي يوضح العلاقة بين المتغيرين النوع وتأييد برنامج تليفزيوني معين.

الفكرة / النوع	موافق جدا	موافق نوعاً ما	لا أدرى	أرفض نوعاً ما	أرفض جداً	المجموع
ذكور	5	37	13	28	5	88
إناث	3	17	8	20	5	53
المجموع	8	54	21	48	10	141

والمطلوب حساب قيمة كا^2 مع بيان مدى دلالتها إحصائيا عند مستوى دلالة 0.05؟

الحل:

حساب التكرار المتوقع ($ت_م$):

$ت_م$ للخلية الأولى (5)= $\frac{8 \times 88}{141} = 5$

$ت_م$ للخلية الثانية (37)= $\frac{54 \times 88}{141} = 33.7$

$ت_م$ للخلية الثالثة (13)= $\frac{21 \times 88}{141} = 13.1$

$ت_م$ للخلية الرابعة (28)= $\frac{48 \times 88}{141} = 29.95$

$ت_م$ للخلية الخامسة (5)= $\frac{10 \times 88}{141} = 6.24$

$ت_م$ للخلية السادسة (3)= $\frac{8 \times 53}{141} = 3$

$ت_م$ للخلية السابعة (17)= $\frac{54 \times 53}{141} = 20.29$

$ت_م$ للخلية الثامنة (8)= $\frac{21 \times 53}{141} = 7.89$

$ت_م$ للخلية التاسعة (28)= $\frac{48 \times 53}{141} = 18$

$ت_م$ للخلية العاشرة (5)= $\frac{10 \times 53}{141} = 3.75$

حساب كا² المحسوبة:

نكون الجدول التالي:

$ت_و$	$ت_م$	$ت_و - ت_م$	$(ت_و - ت_م)^2$	$\frac{(ت_و - ت_م)^2}{ت_م}$
5	37	5	0	0
37	33.7	3.3	10.9	0.32

13	13.1	0.1-	0.01	0
28	29.95	1.59-	3.8	0.13
5	6.24	1.24-	1.5	0.24
3	3	0	0	0
17	20.29	3.29-	10.8	0.53
8	7.89	0.11	0.01	0
20	18	2	4	0.22
5	3.75	1.25	1.56	0.42
–	–	–	مجموع	1.86

من الجدول مباشرة فان مجموع العمود الأخير يعطينا قيمة كا^2

كا^2 المحسوبة = 1.86.

حساب كا^2 الجدولية:

لحسابها يلزم حساب درجة الحرية ومستوى الدلالة:

درجة الحرية = (عدد الصفوف – 1) × (عدد الأعمدة – 1)

= (2 – 1) × (5 – 1) = 1 × 4 = 4

مستوى الدلالة = 0.05.

بالبحث فى جداول كا^2 عند درجة حرية = 4 ومستوى دلالة 0.05 نجد قيمة كا^2 الجدولية = 9.488

تحديد مدى دلالة كا^2:

نقارن قيمة كا^2 المحسوبة بقيمة كا^2 الجدولية نجد أن:

قيمة كا^2 المحسوبة = 1.86 < قيمة كا^2 الجدولية = 9.488

لذا فان كا^2 ليست دالة إحصائية عند مستوى دلالة 0.05.

7 – الحالة السابعة: حساب كا^2 لدلالة فروق النسب المرتبطة

نحسب قيمة كا^2 لدلالة فروق النسب المرتبطة بالجدول الرباعى الخلايا 2×2 من العلاقة:

$$\text{كا}^2 = \frac{(\text{ب} - \text{جـ})^2}{\text{ب} + \text{جـ}}$$

حيث أن ب، جـ هم خلايا بالجدول الرباعى كما بالشكل:

أ	ب
جـ	د

مثال:

احسب قيمة كا2 لدلالة فروق النسب المرتبطة التالية مع بيان مدى دلالتها إحصائيا عند مستوى دلالة 0.05.

النوع / الفكرة	ذكور	إناث	مج
مؤيد	25	15	40
معارض	5	55	60
مج	30	70	100

الحل:

حساب قيمة كا^2 المحسوبة:

$$كا^2 = \frac{(5 - 15)^2}{5 + 15}$$

حساب كا^2 الجدولية:

لحسابها يلزم حساب درجة الحرية ومستوى الدلالة:

درجة الحرية = (عدد الصفوف – 1) × (عدد الأعمدة – 1)

$= (1 - 2) \times (1 - 2) = 1 \times 1 = 1$

مستوى الدلالة = 0.05.

بالبحث في جداول كا^2 عند درجة حرية = 1 ومستوى دلالة 0.05 نجد قيمة كا^2 الجدولية = 3.841

تحديد مدى دلالة كا^2:

نقارن قيمة كا^2 المحسوبة بقيمة كا^2 الجدولية نجد أن:

قيمة كا^2 المحسوبة = 5 > قيمة كا^2 الجدولية = 3.841

لذا فان كا^2 دالة إحصائية عند مستوى دلالة 0.05.

تمارين

1 - من الجدول الرباعي التالي:

ص \ س	نعم	لا	مج
مؤيد	25	15	40
معارض	23	27	50
مج	48	42	90

احسب قيمة $كا^2$ في كل من الحالات التالية:

- بالقانون العام
- بالطريقة المختصرة

ثم بين مدى دلالتها إحصائيا عند مستوى دلالة 0.05.

2 - احسب كا2 من الجدول التالي:

ثم بين مدى دلالتها إحصائيا عند مستوى 0.05

ذكور	4	6	8	3	6
إناث	2	10	25	9	14

3 – من الجدول التالي:

الجنس / الإجابة	ذكور	إناث	مج
موافق	32	22	54
معارض	14	10	24
محايد	4	8	12
مج	50	40	90

احسب قيمة كا²

ثم بين مدى دلالتها إحصائيا عند مستوى دلالة 0.05.

4 – احسب كا² لدلالة فروق النسب المرتبطة التالية مع بيان دلالتها الإحصائية.

30	10
23	27

الفصل العاشر

معاملات الارتباط - الانحدار

أولاً: الارتباط ومعناه.

ثانياً: أنواع الارتباط.

ثالثاً: معامل الاقتران.

رابعاً: معامل فاى.

خامساً: معامل التوافق.

سادساً: معامل ارتباط بيرسون.

سابعاً: معامل ارتباط الرتب لسبيرمان.

ثامناً: معنى الانحدار.

تاسعاً: معادلة خط انحدار ص/س.

عاشراً: معادلة خط انحدار س/ص.

الارتباط ومعناه

تركز عدد من البحوث الاجتماعية على تحليل العلاقة بين أكثر من متغير حيث يهتم الباحث بتحديد كيف وإلى أي مدى يرتبط متغيرات أو أكثر، والإحصاءات المستخدمة في التحليلات ثنائية المتغير، فالمنطق متشابه إلى حد كبير وإن كانت الإحصاءات المستخدمة في دراسة العلاقات متعددة المتغير تتسم بدرجة كبيرة من التعقيد.

وعند تحليل العلاقة بين متغيرين يهتم الباحث بالإجابة عن ثلاثة تساؤلات هل ترتبط هذه المتغيرات؟ وما هو اتجاه وشكل الارتباط الموجود؟ هل هناك احتمال أن يكون الارتباط الذي تمت ملاحظته بين حالات العينة أحد خصائص المجتمع البحثي أم أن هذا الارتباط هو نتاج لصغر حجم العينة التي قد تكون غير ممثلة للمجتمع البحثي؟

يمكن تحديد الارتباط بين متغيرين من خلال استخدام مجموعة من الإحصاءات تعرف باسم معاملات الارتباط ومعامل الارتباط هو رقم يلخص التحسن في تخمين القيم على متغير واحد لأي حالة على أساس معرفة قيم المتغير الثاني، فكلما ارتفع المعامل قوي الارتباط، ومن ثم تحسنت قدرتنا التنبؤية أو التفسيرية. وتتراوح معاملات الارتباط بين صفر وواحد (أو – 1)، وتشير القيم التي تقترب من 1 إلى وجود ارتباط قوي نسبياً أما تلك التي تقترب من صفر فتشير إلى ارتباط ضعيف نسبياً. ويتطلب كل مستوى قياس أنواع مختلفة من الحسابات وبالتالي فلكل من هذه المستويات اختبارات ارتباط مختلفة.

إضافة إلى حجم الارتباط يهتم الباحث بمعرفة اتجاه العلاقة بين المتغيرين فهل هي علاقة طردية أو عكسية، وتجدر الإشارة هنا إلى أن مفهوم الاتجاه ليس له معنى على مستوى القياس الأسمي، حيث إن الأرقام على هذا المستوى من القياس مجرد عناوين للفئات،

وبالتالي لا تتغير إشارات معاملات الارتباط الاسمية فكلها موجبة وتشير إلى مدى قوة الارتباط، أما على مستوى قياس الفترة فإن الإشارات تتغير ولها دلالات هندسية على درجة عالية نسبياً من التعقيد.

وأخيراً يهتم الباحث باختبارات الدلالة الإحصائية وهي الاختبارات التي توضح احتمال أن تكون العلاقات التي يلاحظها الباحث نتاج التحيز في عملية الاختبار بدلاً من أن تعكس علاقات موجودة فعلاً في مجمع البحث.

أنواع الارتباط:

بالطبع عرفنا أن قيمة معامل الارتباط محصورة فى الفترة المغلقة] 1-، 1 [وتتحدد نوعية الارتباط من الجدول التالى:

قيمة معامل الارتباط	نوع الارتباط
1+	ارتباط طردى تام
من 0.7 إلى أقل من 1+	ارتباط طردى قوى
من 0.4 إلى أقل من 0.7	ارتباط طردى متوسط
من صفر إلى أقل من 0.4	ارتباط طردى ضعيف
صفر	الارتباط منعدم
1-	ارتباط عكسي تام
من -0.7 إلى أقل من -1	ارتباط عكسى قوى
من -0.04 إلى أقل من -0.7	ارتباط عكسى متوسط
من صفر إلى أقل من -.04	ارتباط عكسى ضعيف

طرق حساب الارتباط

1 - معامل الاقتران:

يستخدم معامل الاقتران لحساب قيمة معامل الارتباط عندما يكون المتغيران المراد قياس الارتباط بينهم صفات والجدول المزدوج الذي يمثل العلاقة بينهم مكون من (4) خلايا فقط دون خلايا المجموع نستخدم القانون التالى لمعامل الاقتران:

$$\text{معامل الاقتران} = \frac{\text{أ} \times \text{د} - \text{ب} \times \text{جـ}}{\text{أ} \times \text{د} + \text{ب} \times \text{جـ}}$$

حيث أ، ب، جـ، د هم الخلايا الأربع للجدول رباعى الخلايا كما بالشكل:

ب	أ
د	جـ

مثال:

قام أحد الباحثين بعمل بحث عن نسب المدخنيين من النوعين الذكور والإناث فحصل على بيانات الجدول التالى:

مج	إناث	ذكور	النوع / التدخين
40	15	25	يدخن
60	55	5	لا يدخن
100	70	30	مج

والمطلوب حساب قيمة معامل الارتباط بالطريقة المناسبة مع بيان نوع هذا الارتباط؟

الحل:

الجدول مكون من أربعة خلايا فقط والمتغيران صفات لذا نستخدم معامل الاقتران:

$$\text{معامل الاقتران} = \frac{\text{أ} \times \text{د} - \text{ب} \times \text{جـ}}{\text{أ} \times \text{د} + \text{ب} \times \text{جـ}}$$

$$\text{معامل الاقتران} = \frac{5 \times 15 - 55 \times 25}{5 \times 15 + 55 \times 25} = \frac{1300}{1450}$$

معامل الاقتران = 0.89

تحديد نوع الارتباط:

ارتباط طردى قوى.

2 - معامل فاى:

يستخدم معامل فاى لحساب قيمة معامل الارتباط عندما يكون المتغيران المراد قياس الارتباط بينهم صفات أيضاً والجدول المزدوج الذي يمثل العلاقة بينهم مكون من (4) خلايا فقط دون خلايا المجموع نستخدم القانون التالى لحساب لمعامل فاى:

$$\text{معامل فاى} = \frac{\text{أ} \times \text{د} - \text{ب} \times \text{جـ}}{\sqrt{\text{هـ} \times \text{و} \times \text{ز} \times \text{ح}}}$$

حيث أ، ب، جـ، د، هـ، و، ز، ح

هم خلايا الجدول الرباعى الخلايا كما بالشكل التالى:

النوع / الفكرة	ذكور	إناث	المجموع
مؤيد	أ	ب	ح
معارض	جـ	د	ز
المجموع	هـ	و	ن

والسؤال الآن: متى يستخدم معامل الاقتران ومتى يستخدم معامل فاى رغم تشابهها فى الشروط؟

يستخدم معامل فاى إذا كنا نريد استخدام جميع خلايا الجدول أو إذا كنا نريد الحصول على القيمة الأقل لمعامل الارتباط أو الأدق أما بخلاف ذلك نستخدم معمل الاقتران.

مثال:

قام أحد الباحثين بعمل بحث عن نسب المدخنين من النوعين الذكور والإناث فحصل على بيانات الجدول التالى:

النوع / التدخين	ذكور	إناث	مج
يدخن	25	15	40
لا يدخن	5	55	60
مج	30	70	100

والمطلوب حساب قيمة معامل الارتباط بالطريقة المناسبة للحصول على القيمة الأقل والأدق لمعامل الارتباط مع بيان نوع هذا الارتباط؟

الحل:

الجدول مكون من أربعة خلايا فقط والمتغيران صفات والمطلوب الحصول على القيمة الأقل والأدق لمعامل الارتباط لذا نستخدم معامل فاى:

$$\text{معامل فاى} = \frac{\text{أ} \times \text{د} - \text{ب} \times \text{جـ}}{\sqrt{\text{هـ} \times \text{و} \times \text{ز} \times \text{ح}}}$$

$$\text{معامل فاى} = \frac{5 \times 15 - 55 \times 25}{40 \times 60 + 70 \times 30} = \frac{1300}{2245}$$

معامل فاى = 0.58

تحديد نوع الارتباط:

ارتباط طردي متوسط.

التعليق:

نلاحظ أن قيمة معامل الاقتران أكبر من قيمة معامل فاى لحساب قيمة الارتباط لنفس المثال حيث أن معامل فاى أدق من معامل الاقتران لأنه يستخدم جميع خلايا الجدول.

3 - معامل التوافق:

يستخدم معامل التوافق لحساب قيمة معامل الارتباط عندما يكون المتغيران المراد قياس الارتباط بينهم صفات أيضاً والجدول المزدوج الذي يمثل العلاقة بينهم يزيد عدد خلاياه عن (4) خلايا دون خلايا المجموع ونستخدم القانون التالى لحساب قيمة معامل التوافق:

$$\text{معامل التوافق} = \sqrt{\frac{\text{جـ} - 1}{\text{جـ}}}$$

حيث تحسب (جـ) من العلاقة:

$$\text{جـ} = \text{مجـ} \frac{\text{مربع الخلية}}{\text{مجموع صف الخلية} \times \text{مجموع عمود الخلية}}$$

مثال:

قام أحد الباحثين بعمل بحث عن المدخنين ومدى تأثرهم بمشاهدة برنامج خمسة لصحتك فحصل على بيانات الجدول التالى:

التدخين / مشاهدة البرنامج	يدخن	لا يدخن	مج
دائماً يشاهد البرنامج	62	116	178
غالباً يشاهد البرنامج	17	176	193
أحياناً يشاهد البرنامج	5	73	78
لا يشاهد البرنامج	3	20	23
مج	87	385	472

والمطلوب حساب قيمة معامل الارتباط بالطريقة المناسبة مع بيان نوع هذا الارتباط؟

الحل:

الجدول تزيد عدد خلاياه عن أربعة خلايا والمتغيران صفات لذا نستخدم معامل التوافق:

$$\text{معامل التوافق} = \sqrt{\frac{\text{جـ} - 1}{\text{جـ}}}$$

حيث تحسب (جـ) من العلاقة:

$$\text{جـ} = \text{مجـ} \frac{\text{مربع الخلية}}{\text{مجموع صف الخلية} \times \text{مجموع عمود الخلية}}$$

$$جـ = \frac{(62)^2}{87 \times 178} + \frac{(17)^2}{87 \times 193} + \frac{(5)^2}{87 \times 78} + \frac{(3)^2}{87 \times 23}$$

$$+ \frac{(116)^2}{385 \times 178} + \frac{(176)^2}{385 \times 193} + \frac{(73)^2}{385 \times 78} + \frac{(20)^2}{385 \times 23}$$

جـ = 0.248 + 0.017 + 0.004 + 0.005 + 0.196 + 0.417 + 0.178 + 0.045 = 1.11

$$\text{معامل التوافق} = \sqrt{\frac{1 - 1.11}{1.11}}$$

معامل التوافق = 0.32

تحديد نوع الارتباط:

ارتباط طردي ضعيف.

4 - معامل ارتباط بيرسون:

يستخدم معامل ارتباط بيرسون لحساب قيمة معامل الارتباط عندما يكون المتغيران المراد قياس الارتباط بينهم متغيرات كمية ويشترط تساوى عدد حالات كلاً من المتغيرين ونستخدم القانون التالى لحساب قيمة معامل ارتباط بيرسون:

ر: هو معامل ارتباط بيرسون ويحسب من العلاقة:

$$ر = \frac{\text{ن مجـ (س×ص) − مجـ س × مجـ ص}}{\sqrt{[\text{ن مجـ س}^2 - (\text{مجـ س})^2] \times [\text{ن مجـ ص}^2 - (\text{مجـ ص})^2]}}$$

مثال:

الجدول التالى يوضح درجات مجموعة من الطلاب فى اختبار تم إجراؤه على نفس الطلاب مرتين متتاليتين والمطلوب حساب قيمة معامل الارتباط لبيرسون بين درجات الاختبارين؟

درجة الاختبار الأول	3	5	9	8	2
درجة الاختبار الأول	4	6	7	4	3

الحل:

نفترض أن درجات الاختبار الأول هى "س" ودرجات الاختبار الثانى هى "ص"

ثم نكون الجدول التالى:

س	ص	س×ص	$س^2$	$ص^2$
3	4	12	9	16
5	6	30	25	36
9	7	63	81	49
8	4	32	64	16
2	3	6	4	9
27	24	143	183	126

حساب معامل الارتباط لبيرسون:

$$ر = \frac{ن\ مجـ(س×ص) - مجـ س × مجـ ص}{\sqrt{[ن\ مجـ س^2 - (مجـ س)^2] × [ن\ مجـ ص^2 - (مجـ ص)^2]}}$$

نعوض فى المعادلة السابقة:

$$ر = \frac{5 × 143 - 27 × 24}{\sqrt{[5 × 183 - (27)^2] × [5 × 126 - (24)^2]}}$$

ر = 0.668

تحديد نوع الارتباط:

ارتباط طردي متوسط.

5 - معامل ارتباط الرتب لسبيرمان:

يستخدم معامل ارتباط الرتب لسبيرمان لحساب قيمة معامل الارتباط عندما يكون المتغيران المراد قياس الارتباط بينهم متغيرات كمية ويشترط تساوى عدد حالات كلاً من المتغيرين أيضاً ونستخدم القانون التالى لحساب قيمة معامل ارتباط الرتب لسبيرمان:

$$ر = 1 - \frac{6 \text{ مجـ } ف^2}{ن (ن^2 - 1)}$$

حيث:

ر: معامل ارتباط الرتب لسبيرمان

ف = رتب المتغير الأول – رتب المتغير الثانى

ن: عدد الحالات

مثال:

الجدول التالى يوضح درجات مجموعة من الطلاب فى اختبار تم إجراؤه على نفس الطلاب مرتين متتاليتين والمطلوب حساب قيمة معامل ارتباط الرتب لسبيرمان بين درجات الاختبارين؟

درجة الاختبار الأول	3	5	9	8	2
درجة الاختبار الأول	4	6	7	4	3

الحل:

نفترض أن درجات الاختبار الأول هى "س" ودرجات الاختبار الثانى هى "ص"

ثم نكون الجدول التالى:

مع ملاحظة أنه إذا تم ترتيب قيم س تصاعدي لابد من ترتيب قيم ص تصاعدي والعكس بالعكس.

وهنا سوف نرتب القيم تصاعدي.

مع ملاحظة أنه إذا تساوى عددان أو أكثر فى القيمة يأخذ كل منهم متوسط ترتيبهم.

فمثلاً المتغير ص يوجد به رقمان متساويان هما (4،4) وترتبهما (2،3) إذا يأخذ كل منهم متوسط الترتيب (3+2)/ 2 = 5 /2 = 2.5.

س	ص	رتب س	رتب ص	ف	$ف^2$
3	4	2	2.5	0.5-	0.25
5	6	3	4	1-	1
9	7	5	5	0	0
8	4	4	2.5	1.5	2.25
2	3	1	1	0	0
مج					3.5

حساب معامل ارتباط الرتب لسبيرمان:

$$ر = 1 - \frac{6 \text{ مجـ } ف^2}{ن\,(ن^2 - 1)}$$

$$ر = 1 - \frac{6 \times 3.5}{5\,(25 - 1)}$$

$$ر = 1 - \frac{21}{24 \times 5}$$

$ر = 1 - 0.175 = 0.825$

تحديد نوع الارتباط:

ارتباط طردي قوى.

معنى الانحدار:

يهدف الانحدار إلى الإفادة من الارتباط فى التنبؤ، فإذا علمنا معامل ارتباط درجات اختبار الحساب بدرجات اختبار الجبر، وعلمنا درجة أى طالب فى اختبار الحساب فإننا نستطيع أن نتنبأ بدرجته فى الجبر وإذا علمنا درجة أى طالب آخر فى اختبار الجبر فإننا نستطيع أن نتنبأ بدرجته فى الحساب.

وقد سمى هذا المفهوم الإحصائي بالانحدار لأنه ينحدر فى تقديره الدرجات المختلفة نحو المتوسط ولذا تسمى معادلات الانحدار أحياناً بمعادلات خطوط المتوسطات.

حساب الانحدار:

تعتمد معادلات الانحدار معاملات الارتباط وعلى الانحرافات المعيارية وعلى المتوسطات فهى بذلك تستعين بأهم المقاييس الإحصائية فى حسابها لهذا التنبؤ.

أولاً: معادلة خط انحدار ص/س:

تتلخص معادلة خط انحدار ص على س فى الصورة التالية:

$$ص = ر \times \frac{ع\ ص}{ع\ س} (س - م\ س) + م\ س$$

حيث:

ر = معامل ارتباط بيرسون ويحسب من العلاقة:

$$ر = \frac{\text{ن مجـ (س×ص) − مجـ س × مجـ ص}}{\sqrt{[\text{ن مجـ س}^2 − (\text{مجـ س})^2] × [\text{ن مجـ ص}^2 − (\text{مجـ ص})^2]}}$$

ع ص = الانحراف المعياري لقيم ص ويحسب من العلاقة:

$$ع ص = \sqrt{\frac{\text{مجـ ح}^2\text{ ص}}{ن}}$$

ع س = الانحراف المعياري لقيم س ويحسب من العلاقة:

$$ع س = \sqrt{\frac{\text{مجـ ح}^2\text{ س}}{ن}}$$

م س = متوسط قيم المتغير س

م ص = متوسط قيم المتغير ص

مثال:

الجدول التالى يوضح درجات خمس طلاب فى اختبارين الأول س والثانى ص والمطلوب حساب معادلة خط انحدار ص/ س ثم حساب قيمة ص عندما س = 10.

الأفراد	أ	ب	ج	د	هـ
س	2	3	7	18	20
ص	5	7	6	12	10

الحل:

حساب معامل ارتباط بيرسون:

نكون الجدول التالي:

س	ص	س × ص	س²	ص²
2	5	10	4	25
3	7	21	9	49
7	6	42	49	36
18	12	216	324	144
20	10	200	400	100
50	40	489	786	354

$$ر = \frac{ن\ مجـ(س \times ص) - مجـ\ س \times مجـ\ ص}{\sqrt{[ن\ مجـ\ س^2 - (مجـ\ س)^2] \times [ن\ مجـ\ ص^2 - (مجـ\ ص)^2]}}$$

$$ر = \frac{5 \times 489 - 50 \times 40}{\sqrt{[5 \times 786 - (50)^2] \times [5 \times 354 - (40)^2]}}$$

ر = 0.9

حساب المتوسطات:

$$م\ س = \frac{مجـ\ س}{ن} = \frac{50}{5} = 10$$

$$م\ ص = \frac{مجـ\ ص}{ن} = \frac{40}{5} = 8$$

حساب الانحراف المعياري:

نكون الجدول التالي:

س	ص	$ح_س$	$ح^2_س$	$ح_ص$	$ح^2_ص$
2	5	8–	64	3–	9
3	7	7–	49	1–	1
7	6	3–	9	2–	4
18	12	8	64	4	16
20	10	10	100	2	4
			286		34

$$ع_س = \sqrt{\frac{مجـ ح^2_س}{ن}} = \sqrt{\frac{286}{5}} = 7.56$$

$$ع_ص = \sqrt{\frac{مجـ ح^2_ص}{ن}} = \sqrt{\frac{34}{5}} = 2.61$$

حساب معادلة خط انحدار ص/ س:

$$ص = ر \times \frac{ع_ص}{ع_س} (س - م_س) + م_س$$

$$ص = 0.9 \times \frac{2.61}{7.56} (س - 10) + 8$$

$$ص = 0.31 (س - 10) + 8$$

ص = 0.31 س – 3.1 + 8

معادلة خط انحدار ص/ س هى

ص = 0.31 س + 4.9

عندما س = 10 نستطيع التنبؤ بقيمة ص كالتالى:

ص = 0.31 × 10 + 4.9 = 8

ثانياً: معادلة خط انحدار س/ص:

تتلخص معادلة خط انحدار س على ص فى الصورة التالية:

$$س = ر \times \frac{ع\ س}{ع\ ص} (ص - م\ ص) + م\ س$$

حيث:

ر = معامل ارتباط بيرسون ويحسب من العلاقة:

$$ر = \frac{ن\ مجـ (س \times ص) - مجـ\ س \times مجـ\ ص}{\sqrt{[ن\ مجـ\ س^2 - (مجـ\ س)^2] \times [ن\ مجـ\ ص^2 - (مجـ\ ص)^2]}}$$

ع ص = الانحراف المعياري لقيم ص ويحسب من العلاقة

$$ع\ ص = \sqrt{\frac{مجـ\ ح^2\ ص}{ن}}$$

ع س = الانحراف المعياري لقيم س ويحسب من العلاقة:

$$ع\ س = \sqrt{\frac{مجـ\ ح^2\ س}{ن}}$$

م س = متوسط قيم المتغير س

م ص = متوسط قيم المتغير ص

مثال:

الجدول التالى يوضح درجات خمس طلاب فى اختبارين الأول س والثانى ص والمطلوب حساب معادلة خط انحدار س/ ص ثم حساب قيمة س عندما س = 8.

الأفراد	أ	ب	ج	د	هـ
س	2	3	7	18	20
ص	5	7	6	12	10

الحل:

حساب معامل ارتباط بيرسون:

نكون الجدول التالى:

س	ص	س × ص	$س^2$	$ص^2$
2	5	10	4	25
3	7	21	9	49
7	6	42	49	36
18	12	216	324	144
20	10	200	400	100
50	40	489	786	354

$$ر = \frac{ن\ مجـ(س\times ص) - مجـ س \times مجـ ص}{\sqrt{[ن\ مجـ س^2 - (مجـ س)^2] \times [ن\ مجـ ص^2 - (مجـ ص)^2]}}$$

$$ر = \frac{40 \times 50 - 489 \times 5}{\sqrt{[{}^2(40) - 354 \times 5] \times [{}^2(50) - 786 \times 5]}}$$

ر = 0.9

حساب المتوسطات:

$$م\,س = \frac{مجـ\,س}{ن} = \frac{50}{5} = 10$$

$$م\,ص = \frac{مجـ\,ص}{ن} = \frac{40}{5} = 8$$

حساب الانحراف المعياري:

نكون الجدول التالي:

س	ص	$ح_س$	$ح^2_س$	$ح_ص$	$ح^2_ص$
2	5	8−	64	3−	9
3	7	7−	49	1−	1
7	6	3−	9	2−	4
18	12	8	64	4	16
20	10	10	100	2	4
			286		34

$$ع\,س = \sqrt{\frac{مجـ\,ح^2\,س}{ن}} = \sqrt{\frac{286}{5}} = 7.56$$

$$ع\,ص = \sqrt{\frac{مجـ\,ح^2\,ص}{ن}} = \sqrt{\frac{34}{5}} = 2.61$$

حساب معادلة خط انحدار س/ ص:

$$س = ر \times \frac{ع س}{ع ص} (ص - م س) + م س$$

$$س = 0.9 \times \frac{7.56}{2.61} (ص - 8) + 10$$

س = 2.6 (ص – 8) + 10

س = 2.6 س – 20.8 + 10

معادلة خط انحدار س/ ص هى

س = 2.6 ص – 10.8

عندما ص = **8** نستطيع التنبؤ بقيمة س كالتالى:

س = 2.6 × 8 – 10.8 = 10

تمارين

1 - احسب قيمة معامل الارتباط بالطريقة المناسبة ثم حدد نوع هذا الارتباط.

الأفراد	أ	ب	جـ	د	هـ
علم النفس س	7	9	14	5	15
الصحة النفسية ص	11	13	15	6	10

2 - احسب قيمة معامل الارتباط بالطريقة المناسبة ثم حدد نوع هذا الارتباط.

الأفراد	1	2	3	4	5	6	7	8	9	10
س	جيد	مقبول	جيد	ممتاز	جيد جدا	مقبول	جيد جدا	جيد جدا	ضعيف جدا	ضعيف
ص	جيد جدا	ممتاز	مقبول	مقبول	مقبول	جيد	جيد	ممتاز	ضعيف	ضعيف جدا

3 - احسب قيمة معامل الارتباط بالطريقة المناسبة ثم حدد نوع هذا الارتباط.

الأفراد	1	2	3	4	5	6	7	8	9	10
درجات الإحصاء س	8	9	11	10	2	4	3	5	7	6
درجات علم النفس ص	9	10	12	11	3	5	4	6	8	7

4 – احسب قيمة معامل الارتباط بالطريقة المناسبة ثم حدد نوع هذا الارتباط.

س	11	12	13	19	7
ص	6	14	8	10	2

5 – من الجدول الرباعي التالي:

س / ص	نعم	لا	مج
مؤيد	25	15	40
معارض	23	27	50
مج	48	42	90

احسب قيمة معامل الارتباط بالطريقة المناسبة ثم حدد نوع الارتباط؟

6 – احسب قيمة معامل الارتباط بالطريقة المناسبة ثم حدد نوع الارتباط؟

ذكور	4	6	8	3	6
إناث	2	10	25	9	14

7 – من الجدول التالي:

الجنس / الإجابة	ذكور	إناث	مج
موافق	32	22	54
معارض	14	10	24
محايد	4	8	12
مج	50	40	90

احسب قيمة معامل الارتباط بالطريقة المناسبة ثم حدد نوع الارتباط؟

8 - احسب قيمة معامل الارتباط بالطريقة الأكثر دقة والأقل قيمة ثم حدد نوع الارتباط؟

30	10
23	27

9 - احسب معادلة خط انحدار ص/ س

ثم احسب قيمة ص عندما س = 10

س	7	9	6	5	3
ص	5	3	6	7	9

10 - احسب معادلة خط انحدار س/ ص

ثم احسب قيمة س عندما ص = 10

س	7	9	6	5	3
ص	5	3	6	7	9

11 - احسب معادلة خط انحدار ص/ س

ثم احسب قيمة ص عندما س = 20

س	21	24	22	23	25
ص	15	12	12	14	11

12 - احسب معادلة خط انحدار س/ ص

ثم احسب قيمة س عندما ص = 10

س	21	24	22	23	25
ص	15	12	12	14	11

13 - احسب معادلة خط انحدار ص/ س

ثم احسب قيمة ص عندما س = 10

س	6	3	4	2	5
ص	8	4	5	3	5

14 - احسب معادلة خط انحدار س/ ص

ثم احسب قيمة س عندما ص = 20

س	6	3	4	2	5
ص	8	4	5	3	5

الفصل الحادي عشر

الثبات والصدق

أولاً: معنى الثبات.

ثانياً: طرق حساب معامل الثبات.

- **طريقة إعادة الاختبار.**
- **طريقة التجزئة النصفية.**

ثالثاً: معنى الصدق.

رابعاً: قياس الصدق.

- **طريقة المقارنة الطرفية.**

معنى الثبات

إذا أجرى اختبار ما على مجموعة من الأفراد ورصدت درجات كل فرد فى هذا الاختبار ثم أعيد إجراء نفس هذا الاختبار على نفس هذه المجموعة ورصدت أيضا درجات كل فرد ودلت النتائج على أن الدرجات التى حصل عليها الطلاب فى المرة الأولى لتطبيق الاختبار هى نفس الدرجات التى حصل عليها هؤلاء الطلاب فى المرة الثانية، نستنتج من ذلك أن نتائج الاختبار ثابتة تماماً لأن نتائج القياس لم تتغير فى المرة الثانية بل ظلت كما كانت قائمة فى المرة الأولى.

حساب الثبات:

حساب معامل الارتباط هو خير طريقة لمقارنة هذه الدرجات التى حصل علها الطلاب فى الاختبارين.

ويحسب معمل الثبات من العلاقة التالية:

$$\text{معامل الثبات} = \frac{2\text{ر}}{1+\text{ر}}$$

حيث:

ر: هو معامل ارتباط بيرسون ويحسب من العلاقة:

$$\text{ر} = \frac{\text{ن مجـ (س×ص)} - \text{مجـ س} \times \text{مجـ ص}}{\sqrt{[\text{ن مجـ س}^2 - (\text{مجـ س})^2] \times [\text{ن مجـ ص}^2 - (\text{مجـ ص})^2]}}$$

طرق حساب معامل الثبات

1 - طريقة إعادة الاختبار:

تقوم فكرة هذه الطريقة على إجراء الاختبار على مجموعة من الأفراد ثم إعادة إجراء نفس الاختبار على نفس مجموعة الأفراد بعد مضى فترة زمنية وهكذا يحصل كل فرد على درجة فى الإجراء الأول للاختبار وعلى درجة أخرى فى الإجراء الثانى للاختبار، وعندما نرصد هذه الدرجات ونحسب معامل ارتباط درجات المرة الأولى بدرجات المرة الثانية فأننا نحصل بذلك على معامل ثبات الاختبار.

مثال:

الجدول التالى يوضح درجات مجموعة من الطلاب فى اختبار تم إجراؤه على نفس الطلاب مرتين والمطلوب حساب قيمة معمل ثبات الاختبار؟

درجة الاختبار الأول	3	5	9	8	2
درجة الاختبار الأول	4	6	7	4	3

الحل:

نفترض أن درجات الاختبار الأول هى "س" ودرجات الاختبار الثانى هى "ص" ثم نكون الجدول التالى:

س	ص	س×ص	$س^2$	$ص^2$
3	4	12	9	16
5	6	30	25	36

9	7	63	81	49
8	4	32	64	16
2	3	6	4	9
27	**24**	**143**	**183**	**126**

حساب معامل الارتباط لبيرسون:

$$\text{ر} = \frac{\text{ن مجـ (س×ص)} - \text{مجـ س} \times \text{مجـ ص}}{\sqrt{[\text{ن مجـ س}^{2} - (\text{مجـ س})^{2}] \times [\text{ن مجـ ص}^{2} - (\text{مجـ ص})^{2}]}}$$

نعوض في المعادلة السابقة:

$$\text{ر} = \frac{24 \times 27 - 143 \times 5}{\sqrt{[^{2}(24) - 126 \times 5] \times [^{2}(27) - 183 \times 5]}}$$

ر = 0.668

$$\text{معامل الثبات} = \frac{0.668 \times 2}{0.668 + 1}$$

معامل الثبات = 0.8

2 - طريقة التجزئة النصفية:

تعتمد هذه الطريقة على تجزئة الاختبار إلى جزأين فقط بحيث يتكون الجزء الأول من الدرجات الفردية للاختبار ويتكون الجزء الثاني من الدرجات الزوجية للاختبار.

مثال:

الأفراد	الأسئلة							
	1	2	3	4	5	6	7	8
1	1	1	1	1	1	0	0	0

2	1	1	1	0	1	1	0	0
3	1	1	0	1	1	0	0	0
4	1	1	1	0	1	1	1	1
5	1	1	1	0	0	1	0	0
6	1	1	1	1	0	0	1	1
7	1	1	1	0	1	1	0	0
8	1	1	1	1	1	1	1	0
9	1	1	1	1	0	0	0	0
10	1	1	1	1	1	1	1	1

الجدول السابق يوضح درجات عشرة طلاب فى اختبار تم تقسيمه إلى ثماني أسئلة والمطلوب حساب قيمة معامل الثبات لدرجات الأسئلة الفردية والزوجية باستخدام طريقة التجزئة النصفية؟

الحل:

نقوم بتجميع درجات الأسئلة الفردية على حدة ونسميها "س" ودرجات الأسئلة الزوجية على حده ونسميها "ص" لكل طالب منفرداً ونضعها فى الجدول التالى:

س الدرجات الفردية	ص الدرجات الزوجية	س×ص	س^2	ص^2
3	2	6	9	4
3	3	9	9	9

4	4	4	2	2
9	16	12	3	4
4	4	4	2	2
9	9	9	3	3
4	9	6	2	3
9	16	12	3	4
4	4	4	2	2
16	16	16	4	4
72	96	82	26	30

حساب معامل الارتباط لبيرسون:

$$\text{ر} = \frac{\text{ن مجـ (س×ص) − مجـ س × مجـ ص}}{\sqrt{[\text{ن مجـ س}^2 - (\text{مجـ س})^2] \times [\text{ن مجـ ص}^2 - (\text{مجـ ص})^2]}}$$

نعوض فى المعادلة السابقة:

$$\text{ر} = \frac{26 \times 30 - 82 \times 10}{\sqrt{[^2(26) - 72 \times 10] \times [^2(30) - 96 \times 10]}}$$

ر = 0.78

$$\text{معامل الثبات} = \frac{0.78 \times 2}{0.78 + 1}$$

معامل الثبات = 0.88

معنى الصدق

الاختبار الصادق يقيس ما وضع لقياسه فاختبار الذكاء الذى يقيس الذكاء فعلاً اختبار صادق مثله فى ذلك كمثل المتر فى قياسه للأطوال والكيلو فى قياسه للأوزان والساعة فى قياسها للزمن وتختلف الاختبارات فى مستويات صدقها تبعاً لاقترابها أو ابتعادها من تقدير تلك الصفة التى تهدف إلى قياسها فاختبار الذكاء الذى يصل فى قياسه لتلك القدرة إلى مستوى 0.8 أصدق فى هذا القياس من أى اختبار آخر للذكاء لا يصل إلى هذا المستوى أى أنه أصدق مثلاً من الاختبار الذى يصل فى قياسه للذكاء إلى مستوى 0.5.

ويحسب مستوى صدق الاختبار بمقارنة نتائجه بنتائج مقياس آخر دقيق لتلك الصفة ويسمى هذا المقياس بالميزان.

قياس الصدق:

طريقة المقارنة الطرفية

تقوم هذه الطريقة على مقارنة متوسط درجات الأقوياء فى الميزان بمتوسط درجات الضعاف فى نفس ذلك الميزان بالنسبة لتوزيع درجات الاختبار ولذا سميت بالمقارنة الطرفية لاعتمادها على الطرف القوى الذى نسميه بأصحاب الميزان القوى والطرف الضعيف الذى نسميه أصحاب الميزان الضعيف.

ولحساب الدلالة الإحصائية للفرق بين أصاحب المستوى القوى والضعيف نستعين بالنسبة الحرجة:

$$ت = \frac{م_2 - م_1}{\sqrt{\frac{ع_1^2}{ن_1} + \frac{ع_2^2}{ن_2}}}$$

حيث:

م1 = متوسط درجات أصحاب الميزان الضعيف

م2 = متوسط درجات أصحاب الميزان القوى

ع12 = تباين درجات أصحاب المستوى الضعيف

ع22 = تباين درجات أصحاب المستوى القوى

ن1 = مجموع تكرارات أصحاب الميزان الضعيف = مجـ ك1

ن2 = مجموع تكرارات أصحاب الميزان القوى = مجـ ك2

ويحسب المتوسط فى البيانات المبوبة من العلاقة:

$$م = \frac{مجـ(س \times ك)}{مجـ ك}$$

حيث «س» هى مركز الفئة وتحسب من العلاقة:

- س = (بداية الفئة الأولى+ نهاية الفئة)/ 2
- ك: هو التكرار

ويحسب التباين من العلاقة:

$$ع^2 = ل^2 \times \left\{ \frac{مجـ(ح^2 \times ك)}{مجـ ك} - \left\{ \frac{مجـ(ح \times ك)}{مجـ ك} \right\}^2 \right\}$$

حيث:

ح = الانحراف ويحسب عن طريق وضع صفر أمام الفئة ذات اكبر تكرار ثم من أسفل (1، 2، 3،......) ومن أعلى (– 1، – 2، – 3،....).

ل = طول الفئة = الفرق بين بدايتي أى فئتين متتاليتين.

تحديد مدى دلالة النسبة الحرجة وصدق الاختبار من عدمه (3)

- إذا كانت النسبة الحرجة < 1.96 يكون الاختبار غير صادق عند مستوى دلالة 0.05.
- 1.96 < النسبة الحرجة < 2.58 يكون الاختبار صادق عن مستوى دلالة 1.96.
- إذا كانت النسبة الحرجة > 2.58 يكون الاختبار صادق عند مستوى دلالة 0.01.

بالطبع المقارنة بالقيمتين (1.96، 2.58) قيم ثابتة لا تتغير.

مثال:

الجدول التالي يوضح العلاقة بين فئات وتكرارات أصحاب مستوى الميزان القوى والضعيف لعدد من الطلاب فى اختبار للذكاء، والمطلوب حساب قيمة معامل الصدق (النسبة الحرجة) وتحديد صدق الاختبار من عدمه عند مستوى دلالة 0.05؟

الفئات	14 – 16	17 – 19	20 – 22	23 – 25	26 – 28	29 – 31
تكرار الميزان الضعيف	4	3	8	0	0	0
تكرار الميزان القوى	0	0	0	5	7	9

الحل:

نكون الجدول التالى:

ف	$ك_1$	$ك_2$	س	$ك_1$×س	$ك_2$×س	ح	ح×$ك_1$	$ح^2$×$ك_1$	ح×$ك_2$	$ح^2$×$ك_2$
14 – 16	4	0	15	60	0	-2	-8	16	0	0
17 – 19	3	0	18	544	0	-1	-3	3	0	0
20 – 22	8	0	21	168	0	0	0	0	0	0

23-25	0	5	24	0	120	1	0	0	5	5
26-28	0	7	27	0	189	2	0	0	14	28
29-31	0	9	30	0	270	3	0	0	27	81
مجموع	15	21	-	282	579	-	11-	19	46	109

حساب المتوسط لأصحاب الميزان الضعيف:

$$\text{م}_1 = \frac{\text{مجـ}(\text{س} \times \text{ك}_1)}{\text{مجـ ك}_1}$$

$$\text{م}_1 = \frac{282}{15} = 18.8$$

حساب المتوسط لأصحاب الميزان القوى:

$$\text{م}_2 = \frac{\text{مجـ}(\text{س} \times \text{ك}_2)}{\text{مجـ ك}_2}$$

$$\text{م}_2 = \frac{282}{15} = 27.5$$

حساب طول الفئة:

ل = الفرقة بين بدايتي أى فئتين متتاليتين

ل = 17 – 14 = 3

حساب التباين لأصحاب الميزان الضعيف:

$$\text{ع}_1^2 = \text{ل}^2 \times \left\{ \frac{\text{مجـ}(\text{ح}^2 \times \text{ك})}{\text{مجـ ك}} - \left\{ \frac{\text{مجـ}(\text{ح} \times \text{ك})}{\text{مجـ ك}} \right\}^2 \right\}$$

$$\text{ع}_1^2 = (3)^2 \times \left\{ \frac{19}{15} - \left\{ \frac{-11}{15} \right\}^2 \right\}$$

$\text{ع}_1^2 = 3.68$

حساب التباين لأصحاب الميزان القوى:

$$\text{ع}_1^2 = \text{ل}^2 \times \left\{ \frac{\text{مجـ}(\text{ح}^2 \times \text{ك})}{\text{مجـ ك}} - \left\{ \frac{\text{مجـ}(\text{ح} \times \text{ك})}{\text{مجـ ك}} \right\}^2 \right\}$$

$$\text{ع}_1^2 = (3)^2 \times \left\{ \frac{109}{21} - \left\{ \frac{46}{21} \right\}^2 \right\}$$

$\text{ع}_1^2 = 33.29$

حساب قيمة ن_1، ن_2:

$\text{ن}_1 = \text{مجـ ك}_1 = 15$

$\text{ن}_2 = \text{مجـ ك}_2 = 21$

حساب قيمة النسبة الحرجة:

$$\text{النسبة الحرجة} = \frac{م_2 - م_1}{\sqrt{\frac{ع_1^2}{ن_1} + \frac{ع_2^2}{ن_2}}}$$

$$\text{النسبة الحرجة} = \frac{18.8 - 27.5}{\sqrt{\frac{3.68}{15} + \frac{33.29}{21}}}$$

النسبة الحرجة = 6.4.

تحديد صدق الاختبار:

قيمة النسبة الحرجة (6.4) > 1.96 عند مستوى دلالة 0.05 لذا فان الاختبار صادق.

تمارين

1 – قمت بتطبيق اختبار على مجموعة من الطلاب فى مادة الإحصاء الاجتماعي مرتين مختلفتين، وحصلت على الدرجات التالية فى الاختبارين.

س	3	3	2	4	2	3	3	4	2	4
ص	2	3	2	3	2	3	2	3	2	4

والمطلوب:

حساب قيمة معامل الثبات بطريقة إعادة الاختبار.

2 – فيما يلى درجات **(5)** طلاب فى اختبارين س، ص.

س	3	4	5	16	22
ص	3	8	6	12	11

والمطلوب:

حساب قيمة معامل الثبات بطريقة إعادة الاختبار.

3 – فيما يلى درجات (5) طلاب فى اختبار تضمن 10 أسئلة:

الأفراد	الأسئلة									
	1	2	3	4	5	6	7	8	9	10
1	4	1	3	2	5	5	4	4	5	3
2	2	3	5	1	4	2	1	1	3	1
3	5	1	2	1	3	1	4	2	5	1
4	2	1	1	2	1	1	1	4	2	1
5	2	2	2	1	1	3	4	1	1	1

والمطلوب:

حساب قيمة معامل الثبات بطريقة التجزئة النصفية.

4 – من الجدول التالى احسب معامل الصدق بطريقة المقارنة الطرفية وبين مدى صدق الاختبار من عدمه.

الفئات	9 – 5	14 – 10	19 – 15	24 – 20	29 – 25	34 – 30
تكرار الميزان الضعيف	2	3	6	4	0	0
تكرار الميزان القوى	0	0	3	7	6	9

5 – من الجدول التالى احسب معامل الصدق بطريقة المقارنة الطرفية وبين مدى صدق الاختبار من عدمه.

الفئات	10 – 6	15 – 11	20 – 16	25 – 21	30 – 26	35 – 31
تكرار الميزان الضعيف	3	6	5	0	0	0
تكرار الميزان القوى	0	0	5	6	7	6


الجداول الإحصائية

جدول كا[2]

درجة الحرية	مستوى الدلالة أو الثقة		
	0.05	0.01	0.001
1	3.84	6.64	10.83
2	5.99	9.21	13.82
3	7.82	11.35	16.27
4	9.49	13.28	18.47
5	11.07	15.09	20.52
6	12.59	16.81	22.46
7	14.07	18.48	24.32
8	15.51	20.09	26.13
9	16.92	21.67	27.88
10	18.31	23.21	29.59
11	19.68	24.73	31.26
12	21.03	26.22	32.91
13	22.36	27.69	34.53
14	23.69	29.14	36.12

15	25.00	30.58	37.70
16	26.30	32.00	39.25
17	27.59	33.41	40.79
18	28.87	34.81	42.31
19	30.14	36.19	43.82
20	31.41	37.57	45.32
21	32.67	38.93	46.80
22	33.92	40.29	48.27
23	35.17	41.64	49.73
24	36.42	42.98	51.18
25	37.65	44.31	52.62
26	38.89	45.64	54.05
27	40.11	46.96	55.48
28	41.34	48.28	56.89
29	42.56	49.59	58.30
30	43.77	50.89	59.70
31	44.99	52.19	61.10
32	46.19	53.49	62.49
33	47.40	54.78	63.87

34	48.60	56.06	65.25
35	49.80	57.34	66.62
36	51.00	58.62	67.99
37	52.19	59.89	69.35
38	53.38	61.16	70.71
39	54.57	62.43	72.06
40	55.76	63.69	73.41
41	56.94	64.95	74.75
42	58.12	66.21	76.09
43	59.30	67.46	77.42
44	60.48	68.71	78.75
45	61.66	69.96	80.08
46	62.83	71.20	81.40
47	64.00	72.44	82.72
48	65.17	73.68	84.03
49	66.34	74.92	85.35
50	67.51	76.15	86.66
51	68.67	77.39	87.97
52	69.83	78.62	89.27

53	70.99	79.84	90.57
54	72.15	81.07	91.88
55	73.31	82.29	93.17
56	74.47	83.52	94.47
57	75.62	84.73	95.75
58	76.78	85.95	97.03
59	77.93	87.17	98.34
60	79.08	88.38	99.62
61	80.23	89.59	100.88
62	81.38	90.80	102.15
63	82.53	92.01	103.46
64	83.68	93.22	104.72
65	84.82	94.42	105.97
66	85.97	95.63	107.26
67	87.11	96.83	108.54
68	88.25	98.03	109.79
69	89.39	99.23	111.06
70	90.53	100.42	112.31
71	91.67	101.62	113.56

72	92.81	102.82	114.84
73	93.95	104.01	116.08
74	95.08	105.20	117.35
75	96.22	106.39	118.60
76	97.35	107.58	119.85
77	98.49	108.77	121.11
78	99.62	109.96	122.36
79	100.75	111.15	123.60
80	101.88	112.33	124.84
81	103.01	113.51	126.09
82	104.14	114.70	127.33
83	105.27	115.88	128.57
84	106.40	117.06	129.80
85	107.52	118.24	131.04
86	108.65	119.41	132.28
87	109.77	120.59	133.51
88	110.90	121.77	134.74
89	112.02	122.94	135.96
90	113.15	124.12	137.19

91	114.27	125.29	138.45
92	115.39	126.46	139.66
93	116.51	127.63	140.90
94	117.63	128.80	142.12
95	118.75	129.97	143.32
96	119.87	131.14	144.55
97	120.99	132.31	145.78
98	122.11	133.47	146.99
99	123.23	134.64	148.21
100	124.34	135.81	149

جدول ت

درجة الحرية	مستوى الدلالة							
طرف واحد	0.1	0.05	0.01	0.005	0.001	0.0005	0.0001	0.00005
طرفين	0.2	0.1	0.05	0.01	0.005	0.001	0.0005	0.0001
2	1.89	2.92	4.30	9.92	14.09	31.60	44.70	100.14
3	1.64	2.35	3.18	5.84	7.45	12.92	16.33	28.01
4	1.53	2.13	2.78	4.60	5.60	8.61	10.31	15.53
5	1.48	2.02	2.57	4.03	4.77	6.87	7.98	11.18
6	1.44	1.94	2.45	3.71	4.32	5.96	6.79	9.08
7	1.41	1.89	2.36	3.50	4.03	5.41	6.08	7.89
8	1.40	1.86	2.31	3.36	3.83	5.04	5.62	7.12
9	1.38	1.83	2.26	3.25	3.69	4.78	5.29	6.59
10	1.37	1.81	2.23	3.17	3.58	4.59	5.05	6.21
11	1.36	1.80	2.20	3.11	3.50	4.44	4.86	5.92
12	1.36	1.78	2.18	3.05	3.43	4.32	4.72	5.70
13	1.35	1.77	2.16	3.01	3.37	4.22	4.60	5.51
14	1.35	1.76	2.14	2.98	3.33	4.14	4.50	5.36
15	1.34	1.75	2.13	2.95	3.29	4.07	4.42	5.24

16	1.34	1.75	2.12	2.92	3.25	4.01	4.35	5.13
17	1.33	1.74	2.11	2.90	3.22	3.97	4.29	5.04
18	1.33	1.73	2.10	2.88	3.20	3.92	4.23	4.97
19	1.33	1.73	2.09	2.86	3.17	3.88	4.19	4.90
20	1.33	1.72	2.09	2.85	3.15	3.85	4.15	4.84
21	1.32	1.72	2.08	2.83	3.14	3.82	4.11	4.78
22	1.32	1.72	2.07	2.82	3.12	3.79	4.08	4.74
23	1.32	1.71	2.07	2.81	3.10	3.77	4.05	4.69
24	1.32	1.71	2.06	2.80	3.09	3.75	4.02	4.65
25	1.32	1.71	2.06	2.79	3.08	3.73	4.00	4.62
26	1.31	1.71	2.06	2.78	3.07	3.71	3.97	4.59
27	1.31	1.70	2.05	2.77	3.06	3.69	3.95	4.56
28	1.31	1.70	2.05	2.76	3.05	3.67	3.93	4.53
29	1.31	1.70	2.05	2.76	3.04	3.66	3.92	4.51
30	1.31	1.70	2.04	2.75	3.03	3.65	3.90	4.48
35	1.31	1.69	2.03	2.72	3.00	3.59	3.84	4.39
40	1.30	1.68	2.02	2.70	2.97	3.55	3.79	4.32
45	1.30	1.68	2.01	2.69	2.95	3.52	3.75	4.27
50	1.30	1.68	2.01	2.68	2.94	3.50	3.72	4.23
55	1.30	1.67	2.00	2.67	2.92	3.48	3.70	4.20

60	1.30	1.67	2.00	2.66	2.91	3.46	3.68	4.17
65	1.29	1.67	2.00	2.65	2.91	3.45	3.66	4.15
70	1.29	1.67	1.99	2.65	2.90	3.43	3.65	4.13
75	1.29	1.67	1.99	2.64	2.89	3.42	3.64	4.11
80	1.29	1.66	1.99	2.64	2.89	3.42	3.63	4.10
85	1.29	1.66	1.99	2.63	2.88	3.41	3.62	4.08
90	1.29	1.66	1.99	2.63	2.88	3.40	3.61	4.07
95	1.29	1.66	1.99	2.63	2.87	3.40	3.60	4.06
100	1.29	1.66	1.98	2.63	2.87	3.39	3.60	4.05
200	1.29	1.65	1.97	2.60	2.84	3.34	3.54	3.97
500	1.28	1.65	1.96	2.59	2.82	3.31	3.50	3.92
1000	1.28	1.65	1.96	2.58	2.81	3.30	3.49	3.91
∞	1.28	1.64	1.96	2.58	2.81	3.29	3.48	3.89

جدول ف

درجة حرية التباين الصغير	درجة حرية التباين الكبير								
	1	2	3	4	5	6	8	12	∞
1	161	200	216	225	230	234	239	244	254
2	18.5	19.0	19.2	19.3	19.3	19.3	19.4	19.4	19.5
3	10.1	9.6	9.3	9.1	9.0	8.9	8.8	8.7	8.5
4	7.7	6.9	6.6	6.4	6.3	6.2	6.0	5.9	5.6
5	6.6	5.8	5.4	5.2	5.1	5.0	4.8	4.7	4.4
6	6.0	5.1	4.8	4.5	4.4	4.3	4.2	4.0	3.7
7	5.6	4.7	4.4	4.1	4.0	3.9	3.7	3.6	3.2
8	5.3	4.5	4.1	3.8	3.7	3.6	3.4	3.3	2.9
9	5.1	4.3	3.9	3.6	3.5	3.4	3.2	3.1	2.7
10	5.0	4.1	3.7	3.5	3.3	3.2	3.1	2.9	2.5
11	4.8	4.0	3.6	3.4	3.2	3.1	3.0	2.8	2.4
12	4.8	3.9	3.5	3.3	3.1	3.0	2.9	2.7	2.3
13	4.7	3.8	3.4	3.2	3.0	2.9	2.8	2.6	2.2
14	4.6	3.7	3.3	3.1	3.0	2.9	2.7	2.5	2.1

15	4.5	3.7	3.3	3.1	2.9	2.8	2.6	2.5	2.1
16	4.5	3.6	3.2	3.0	2.9	2.7	2.6	2.4	2.0
17	4.5	3.6	3.2	3.0	2.8	2.7	2.6	2.4	2.0
18	4.4	3.6	3.2	2.9	2.8	2.7	2.5	2.3	1.9
19	4.4	3.5	3.1	2.9	2.7	2.6	2.5	2.3	1.9
20	4.4	3.5	3.1	2.9	2.7	2.6	2.5	2.3	1.8
21	4.3	3.5	3.1	2.8	2.7	2.6	2.4	2.3	1.8
22	4.3	3.4	3.1	2.8	2.7	2.6	2.4	2.2	1.8
23	4.3	3.4	3.0	2.8	2.6	2.5	2.4	2.2	1.8
24	4.3	3.4	3.0	2.8	2.6	2.5	2.4	2.2	1.7
25	4.2	3.4	3.0	2.8	2.6	2.5	2.3	2.2	1.7
26	4.2	3.4	3.0	2.7	2.6	2.5	2.3	2.2	1.7
27	4.2	3.4	3.0	2.7	2.6	2.5	2.3	2.1	1.7
28	4.2	3.3	3.0	2.7	2.6	2.4	2.3	2.1	1.7
29	4.2	3.3	2.9	2.7	2.5	2.4	2.3	2.1	1.6
30	4.2	3.3	2.9	2.7	2.5	2.4	2.3	2.1	1.6
40	4.1	3.2	2.8	2.6	2.5	2.3	2.2	2.0	1.5
60	4.0	3.2	2.8	2.5	2.4	2.3	2.1	1.9	1.4
120	3.9	3.1	2.7	2.5	2.3	2.2	2.0	1.8	1.3
∞	3.8	3.0	2.6	2.4	2.2	2.1	1.9	1.8	1.0

أهم المراجع

1 - اعتماد علام، يسرى رسلان، أساسيات الإحصاء الاجتماعي، دار الثقافة للنشر والتوزيع.

2 - أنور عطية العدل، مبادئ الإحصاء الاجتماعي، دار المعرفة الجامعية، 1987.

3 - حسن محمد حسن، أساسيات الإحصاء وتطبيقاته، دار المعرفة الجامعية، 1992.

4 - حسن محمد حسن، مبادئ الإحصاء الاجتماعي، دار المعرفة الجامعية، 2000.

5 - خليفة عبد السميع خليفة، الإحصاء التربوي، مكتبة الأنجلو المصرية.

6 - عبد الله عبد الحليم وآخرون، الإحصاء مفاهيم أساسية، 2003.

7 - غريب محمد سيد أحمد، الإحصاء والقياس فى البحث الاجتماعي، دار المعرفة الجامعية، 1989.

8 - غريب محمد سيد أحمد، ناجى بدر إبراهيم، الإحصاء والقياس فى البحث الاجتماعي، دار المعرفة الجامعية، 1997.

9 - فاروق عبد العظيم وآخرون، مبادئ الإحصاء، دار المعرفة الجامعية.

10 - فتحى عبد العزيز أبو راضى، مبادئ الإحصاء الاجتماعي، دار المعرفة الجامعية.

11 - محمد بهجت كشك، مبادئ الإحصاء الاجتماعي، دار المعرفة الجامعية، 1996.

12 - مصطفى زايد، الإحصاء ووصف البيانات، 1989.

13 – http: / /www.mohp.gov.eg /Sec /Heducation /tadrib /5.doc

14 – http: / /www.arab –api.org /course13 /c13_4.htm

15–http: / /dentarab.com /site /index.php?page=show_det&id=178